Applied Numerical Mathematics 180 (2022) 33–54

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Unconditional well-posedness and IMEX improvement of a family of predictor-corrector methods in micromagnetics ☆

Norbert J. Mauser [a], Carl-Martin Pfeiler [b,*], Dirk Praetorius [b], Michele Ruggeri [c]

[a] *Research Platform MMM"Mathematics-Magnetism-Materials" c/o Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria*
[b] *Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040 Vienna, Austria*
[c] *Department of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow G1 1XH, United Kingdom*

ARTICLE INFO

Article history:
Received 1 December 2021
Received in revised form 30 March 2022
Accepted 15 May 2022
Available online 20 May 2022

Keywords:
Landau–Lifshitz–Gilbert equation
Micromagnetism
Finite elements
Implicit-explicit time-marching scheme
Unconditional well-posedness

ABSTRACT

Recently, *Kim & Wilkening (Convergence of a mass-lumped finite element method for the Landau–Lifshitz equation, Quart. Appl. Math., 76, 383–405, 2018)* proposed two novel predictor-corrector methods for the Landau–Lifshitz–Gilbert equation (LLG) in micromagnetics, which models the dynamics of the magnetization in ferromagnetic materials. Both integrators are based on the so-called Landau–Lifshitz form of LLG, use mass-lumped variational formulations discretized by first-order finite elements, and only require the solution of linear systems, despite the nonlinearity of LLG. The first(-order in time) method combines a linear update with an explicit projection of an intermediate approximation onto the unit sphere in order to fulfill the LLG-inherent unit-length constraint at the discrete level. In the second(-order in time) integrator, the projection step is replaced by a linear constraint-preserving variational formulation. In this paper, we extend the analysis of the integrators by proving unconditional well-posedness and by establishing a close connection of the methods with other approaches available in the literature. Moreover, the new analysis also provides a well-posed integrator for the Schrödinger map equation (which is the limit case of LLG for vanishing damping). Finally, we design an implicit-explicit strategy for the treatment of the lower-order field contributions, which significantly reduces the computational cost of the schemes, while preserving their theoretical properties.

☆ **Acknowledgment.** The authors thankfully acknowledge support by the Austrian Science Fund (FWF) through the doctoral school *Dissipation and dispersion in nonlinear PDEs* (grant W1245), the special research program *Taming complexity in partial differential systems* (grant F65), and the project *Reduced order approaches for micromagnetics* (grant P31140); by the Vienna Science and Technology Fund (WWTF) through the project *Schrödinger Equations for QUantum EXperiments (SEQUEX)* (grant MA16-066); and by the University of Vienna research platform MMM ("Mathematics-Magnetism-Materials"). Further, we thank Lukas Exl for fruitful discussions in the early stage of this work.

* Corresponding author.
E-mail addresses: norbert.mauser@univie.ac.at (N.J. Mauser), carl-martin.pfeiler@asc.tuwien.ac.at (C.-M. Pfeiler), dirk.praetorius@asc.tuwien.ac.at (D. Praetorius), michele.ruggeri@strath.ac.uk (M. Ruggeri).

https://doi.org/10.1016/j.apnum.2022.05.008

1. Introduction

1.1. Dynamic micromagnetism

Reliable numerical simulations of magnetic processes occurring at submicrometer length scales are fundamental tools to optimize the design of many technological devices, e.g., magnetic sensors, magnetic logic gates, and hard disk drives. The theoretical background of most simulation packages is the theory of *micromagnetism* [13], a continuum theory which models the magnetic state of a ferromagnetic material at constant temperature in terms of a continuous vector field with constant magnitude, the *magnetization*. A well-accepted model to describe the dynamics of the magnetization is a nonlinear parabolic partial differential equation (PDE) usually referred to as *Landau–Lifshitz–Gilbert equation (LLG)* [33,28], which in the so-called *Landau–Lifshitz (LL) form* reads as

$$\partial_t \boldsymbol{m} = -\frac{1}{1+\alpha^2}\, \boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) - \frac{\alpha}{1+\alpha^2}\, \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})). \tag{1}$$

Here, $\boldsymbol{m}$ denotes the normalized magnetization, which satisfies the nonconvex unit-length constraint $|\boldsymbol{m}| = 1$, $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})$ is the effective field, whose specific expression depends on the Gibbs free energy of the system (see (6) below), and $\alpha \geq 0$ is the Gilbert damping parameter, which incorporates energy dissipation into the model.

Alternative forms of LLG used in the literature, mathematically equivalent to the LL form (1), are the so-called *Gilbert form* of LLG

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) + \alpha\, \boldsymbol{m} \times \partial_t \boldsymbol{m}\,, \tag{2}$$

and

$$\alpha\, \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = -\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})), \tag{3}$$

which we call the *alternative form* of LLG.

The aforementioned need of fast and reliable tools to perform micromagnetic simulations encouraged many works concerned with the numerical analysis of LLG, which will also be the subject of the present paper.

1.2. State of the art

In the last three decades, mathematical questions arising from the micromagnetic theory have been the subject of several studies, from both the analytical and the numerical point of view. For analytical results for LLG, we refer, e.g., to the papers [43,5,29,15,35,22,24,20] and the references therein. For an overview of numerical methods proposed for LLG (up to 2008), we refer to the monograph [39] and the review articles [32,27,17]. More recently, several numerical schemes with a rigorous convergence analysis have been proposed. They differ from each other in the LLG formulation they are based on (usually one among (1)–(3)), in the approach used to impose the unit-length constraint at the discrete level, and in the type of convergence result (plain convergence towards a weak solution of LLG with minimal regularity or convergence with rates towards a sufficiently regular strong solution).

Semi-implicit finite element methods based on (variants of) the LL form (1) of LLG are proposed in [26,7], where *a priori* error estimates, which show their convergence towards a smooth solution of LLG, are also established.

A class of methods referred to as *tangent plane schemes* or *projection methods* [4,11,3,14,1,6,23,21,2] is based on a predictor-corrector approach: At each time-step, first, an update is computed by solving a linear variational problem posed in the discrete tangent space of the current magnetization; second, the update is used to obtain the magnetization at the next time-step. The methods proposed in [4,11,3,14,1,23] are based on a variational formulation of (3) discretized by first-order finite elements to compute an approximation of the linear velocity $\partial_t \boldsymbol{m}$. The magnetization at the next time-step is then obtained via a first-order time-stepping. To impose the unit-length constraint at the vertices of the underlying mesh, the nodal values are projected onto the sphere in [4,11,3,14]. The projection is omitted from the time-stepping in [1,23]: In this case, the approximations do not fulfill the constraint (not even at the vertices of the mesh), but this error can be controlled by the time-step size (in particular, the constraint holds for the solution of LLG towards which the finite element approximation converges). High-order extensions of the tangent plane approach have been proposed in [6,21,2]. The main advantages of this class of methods are that they do not require any time-step restriction for convergence (*unconditional* convergence) [3,14,1,6,23,21] and that, despite the nonlinear nature of LLG, only one linear system per time-step has to be solved.

A numerical scheme based on the Gilbert form (2) of LLG is considered in [10,38]. The method employs mass-lumped first-order finite elements for the spatial discretization and the second-order implicit midpoint rule for the time discretization. The scheme is unconditionally convergent towards a weak solution of LLG, but requires the solution of a nonlinear system of equations per time-step. A similar method, but based on the LL form (1) of LLG, is proposed and analyzed in [18]. The latter approach is motivated by the interest in having an integrator which is robust with respect to the limit cases of (1) in which one of the two terms on the right-hand side tends to zero. Indeed, in the case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \boldsymbol{\Delta m}$, neglecting

the second (dissipative) term on the left-hand side of (1) ($\alpha \to 0$), one obtains the so-called Schrödinger map equation [41], whereas omitting the first (conservative) term, one is led to the harmonic map heat flow [34].

The recent work [31] proposes two predictor-corrector schemes for LLG which aim to combine the features of some of the above integrators. In the first scheme, [31, Algorithm 1], which we denote by `PC1` for the sake of brevity, the predictor is based on the LL form (1) of LLG (like the variational formulation used in [18]) and employs mass-lumping for its discretization (like [10,18]). However, it only requires the solution of one linear system per time-step and uses the nodal projection to impose the unit-length constraint (like the method of [3,14]). The second scheme, [31, Algorithm 2], which we refer to as `PC2`, uses the same predictor as `PC1`, but replaces the nodal projection step with a constraint-preserving mass-lumped (as in [10,18]), but linear (as in [3,14]), variational formulation. In the paper, adapting the proof of [3], the authors show convergence of the approximations generated by `PC1` towards a weak solution of LLG. Moreover, the expected convergence order in time of both methods (first-order for `PC1`, second-order for `PC2`) is empirically verified by means of numerical experiments in 2D.

Note that in the above discussion we have restricted ourselves to methods employing the finite element method for the spatial discretization. For other approaches based on finite differences, we refer, e.g., to [44,19,30,45,16] and the references therein.

1.3. Novelty of the present work

In this work, we improve the theoretical understanding of the predictor-corrector methods proposed in [31].

First, we show that `PC1` is unconditionally well-posed, i.e., for each time-step, the variational problem to be solved admits a unique solution, which is left open in the original paper. By closing this fundamental gap, we show that `PC1` is not only closely related to the first-order tangent plane scheme of [3,14], but actually can even be interpreted as a slight modification of it, which explains why the convergence analysis of the two schemes is almost identical. Furthermore, following [14], we propose an implicit-explicit (IMEX) version of `PC1`. When considering magnetization dynamics involving the full effective field—more precisely, dynamics including the nonlocal stray field—the proposed adaptation is computationally much more attractive: The IMEX version `PC1+IMEX` avoids the costly inner iteration in the solver of the original scheme, while preserving the experimental first-order accuracy of `PC1`, which we confirm by numerical studies in 3D.

Second, we consider the analysis of `PC2`. While the conservation of the unit-length constraint at the vertices of the mesh in `PC1` is guaranteed (at machine precision) also in practical computations (since it is directly enforced in the method using the nodal projection), the one guaranteed by `PC2`, which follows from the variational formulation of the corrector, is lost in practice due to the inevitable use of inexact (iterative) solvers for the solution of the arising linear systems. Hence, although the predictors of `PC1` and `PC2` coincide in theory, the well-posedness analysis of (the predictor of) `PC1` does not transfer to a practical version of `PC2`. To cope with this problem, we establish a decomposition of the finite element space, which does not only allow us to prove unconditional well-posedness of the practical version of `PC2`, but also to extend the result, for both `PC1` and `PC2` (theoretical and practical), to the limit case $\alpha = 0$ (Schrödinger map equation). Moreover, following [38,21], we adopt the IMEX treatment also for `PC2`. In particular, in the presence of the nonlocal stray field, the proposed method `PC2+IMEX` is computationally much more attractive than its fully implicit counterpart `PC2`, while conserving the experimental second-order accuracy in time. Again, these claims are confirmed in our numerical studies. Stability and convergence of `PC2`, not addressed in [31], remain open also in our analysis and will be the subject of future research. In this paper, we shed some light on this question by means of some surprising numerical experiments.

1.4. Outline

We conclude this section by collecting some general notation and basic vector identities used throughout the work (Section 1.5). In Section 2, we formulate the initial boundary value problem for LLG in which we are interested, we recall the notion of a weak solution and introduce the basic ingredients of the discretization. Section 3 is devoted to the first-order method: After proving unconditional well-posedness of `PC1` in Section 3.2, we propose an IMEX adaptation (Section 3.3) overcoming the inefficiency drawbacks of the original version, while preserving unconditional well-posedness, stability, convergence (Section 3.4), and accuracy. Section 4 is devoted to the second-order method: In Section 4.2, we first prove unconditional well-posedness of `PC2`. Subsequently, in Section 4.3, we extend the unconditional well-posedness result to the more general formulation of the second-order algorithm, where discrete unit-length of the iterates is not assumed. This covers, in particular, the practical version of the scheme incorporating the inevitable use of inexact (iterative) linear solvers. Section 4.4 closes with a second-order accuracy preserving IMEX modification overcoming the inefficiency drawbacks of `PC2`. Section 5 provides numerical studies validating the applicability (Section 5.1) and the expected accuracy (Section 5.2) of the IMEX integrators proposed in this work. Finally, in Section 5.3, we numerically investigate the stability of `PC2`.

1.5. General notation and vector identities

Throughout this work, we use the standard notation for Lebesgue, Sobolev, and Bochner spaces and norms. Vector-valued functions are indicated by bold letters. Bold letters are also used for vector-valued and matrix-valued function spaces, e.g., both $L^2(\Omega;\mathbb{R}^3)$ and $L^2(\Omega;\mathbb{R}^{3\times 3})$ are denoted by $\boldsymbol{L}^2(\Omega)$. We denote by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ the scalar product and the norm of

$\boldsymbol{L}^2(\Omega)$, respectively, while $|\cdot|$ denotes the Euclidean norm of a vector in $\mathbb{R}^3$ or the Frobenius norm of a matrix in $\mathbb{R}^{3\times 3}$. To abbreviate notation in proofs, we write $A \lesssim B$ when $A \le cB$ for some generic constant $c > 0$, which is clear from the context and always independent of the discretization parameters. For vector-valued functions $\boldsymbol{f}, \boldsymbol{g}\colon \Omega \to \mathbb{R}^3$ we use the notation

$$-\boldsymbol{g} \times \nabla \boldsymbol{f} := \nabla \boldsymbol{f} \times \boldsymbol{g} := (\partial_1 \boldsymbol{f} \times \boldsymbol{g}, \partial_2 \boldsymbol{f} \times \boldsymbol{g}, \partial_3 \boldsymbol{f} \times \boldsymbol{g})\colon \Omega \to \mathbb{R}^{3\times 3}.$$

We conclude this section by recalling five vector identities used regularly in this work

$$\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}, \tag{4a}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{a} = 0, \tag{4b}$$

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\, \boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\, \boldsymbol{c}, \tag{4c}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}), \tag{4d}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a} \cdot \boldsymbol{c})\, (\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{b} \cdot \boldsymbol{c})\, (\boldsymbol{a} \cdot \boldsymbol{d}), \tag{4e}$$

which hold true for arbitrary $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d} \in \mathbb{R}^3$.

2. Problem formulation

2.1. Landau–Lifshitz–Gilbert equation

Given a bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$ and $T > 0$, we define the space-time cylinder $\Omega_T := \Omega \times (0, T)$. We consider the following initial boundary value problem

$$(1 + \alpha^2)\, \partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) - \alpha\, \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})) \qquad \text{in } \Omega_T, \tag{5a}$$

$$\partial_{\boldsymbol{n}} \boldsymbol{m} = \boldsymbol{0} \qquad \text{on } \partial\Omega \times (0, T), \tag{5b}$$

$$\boldsymbol{m}(0) = \boldsymbol{m}^0 \qquad \text{in } \Omega. \tag{5c}$$

The unknown is the normalized magnetization $\boldsymbol{m}\colon \Omega_T \to \mathbb{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3\colon |\boldsymbol{x}| = 1\}$. In (5a), the effective field

$$\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2\, \Delta \boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f} \tag{5d}$$

is the negative functional derivative of the Gibbs free energy

$$\mathcal{E}(\boldsymbol{m}) = \frac{\ell_{\mathrm{ex}}^2}{2} \int_\Omega |\nabla \boldsymbol{m}|^2 \,\mathrm{d}\boldsymbol{x} - \frac{1}{2} \int_\Omega \boldsymbol{\pi}(\boldsymbol{m}) \cdot \boldsymbol{m} \,\mathrm{d}\boldsymbol{x} - \int_\Omega \boldsymbol{f} \cdot \boldsymbol{m} \,\mathrm{d}\boldsymbol{x}, \tag{6}$$

where $\ell_{\mathrm{ex}} > 0$ is the exchange length, $\boldsymbol{\pi}\colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ is a linear, continuous, and self-adjoint operator which collects all lower-order contributions such as uniaxial magnetocrystalline anisotropy and the nonlocal stray field, and $\boldsymbol{f}\colon \Omega_T \to \mathbb{R}^3$ is the applied external field. The equation is supplemented with homogeneous Neumann boundary conditions (5b) and the initial condition (5c), where $\boldsymbol{m}^0\colon \Omega \to \mathbb{S}^2$ denotes a given initial state.

Taking the scalar product of (5a) with $\boldsymbol{m}$, (4b) yields that $0 = \partial_t \boldsymbol{m} \cdot \boldsymbol{m}$ in Ω_T. Since $|\boldsymbol{m}^0| = 1$ in Ω by assumption and $\partial_t(|\boldsymbol{m}|^2/2) = \partial_t \boldsymbol{m} \cdot \boldsymbol{m} = 0$, it follows that $|\boldsymbol{m}| = 1$ in Ω_T. Moreover, any solution of (5a) satisfies the energy law

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(\boldsymbol{m}, \boldsymbol{f}) = -\alpha \int_\Omega |\partial_t \boldsymbol{m}|^2 \,\mathrm{d}\boldsymbol{x} - \int_\Omega \partial_t \boldsymbol{f} \cdot \boldsymbol{m} \,\mathrm{d}\boldsymbol{x}. \tag{7}$$

From this, we see that the Gilbert damping constant α modulates the dissipation of the system. In particular, if $\alpha = 0$ and $\boldsymbol{f}$ is constant in time, then the energy is conserved. The PDE inherent constraint $|\boldsymbol{m}| = 1$ in Ω_T and the energy law (7) should be satisfied (at the discrete level) by any feasible numerical method.

2.2. Weak solution

We recall the notion of a weak solution of (5), which extends the one introduced in [5].

Definition 2.1. Let $\boldsymbol{m}^0 \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ and $\boldsymbol{f} \in C^1([0, T]; \boldsymbol{L}^2(\Omega))$. A vector field $\boldsymbol{m}\colon \Omega_T \to \mathbb{R}$ is called a weak solution of (5), if the following properties are satisfied:

(i) $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T) \cap L^\infty(0, T; \boldsymbol{H}^1(\Omega))$ with $|\boldsymbol{m}| = 1$ a.e. in Ω_T;
(ii) $\boldsymbol{m}(0) = \boldsymbol{m}^0$ in the sense of traces;

(iii) for all $\boldsymbol{w} \in \boldsymbol{H}^1(\Omega_T)$, it holds that

$$\begin{aligned}
&\int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{w}(t)\rangle \,\mathrm{d}t - \alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{w}(t)\rangle \,\mathrm{d}t \\
&\quad = \ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{m}(t) \times \nabla \boldsymbol{m}(t), \nabla \boldsymbol{w}(t)\rangle \,\mathrm{d}t - \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{\pi}(\boldsymbol{m}(t)), \boldsymbol{w}(t)\rangle \,\mathrm{d}t \\
&\qquad - \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{f}(t), \boldsymbol{w}(t)\rangle \,\mathrm{d}t;
\end{aligned} \tag{8}$$

(iv) it holds that

$$\mathcal{E}(\boldsymbol{m}(T)) + \alpha \int_0^T \|\partial_t \boldsymbol{m}(t)\|^2 \,\mathrm{d}t + \int_0^T \langle \partial_t \boldsymbol{f}(t), \boldsymbol{m}(t)\rangle \,\mathrm{d}t \le \mathcal{E}(\boldsymbol{m}^0). \tag{9}$$

We note that (8) is a variational formulation in space-time of the Gilbert form (2) of LLG, and that (9) is a weaker version of the energy law (7).

2.3. Discretization

For the temporal discretization, given $L \in \mathbb{N}$, we consider a partition $\{t_\ell\}_{\ell=0,\dots,L}$ of the time interval $[0,T]$ with uniform time-step size $k := T/L > 0$, i.e., $t_\ell = \ell k$ for all $\ell = 0,\dots,L$. Given a finite sequence of functions $\{\boldsymbol{u}^\ell\}_{\ell=0,\dots,L}$, we define

$$\boldsymbol{u}^{\ell+1/2} := \frac{\boldsymbol{u}^{\ell+1} + \boldsymbol{u}^\ell}{2} \quad \text{and} \quad d_t \boldsymbol{u}^{\ell+1} := \frac{\boldsymbol{u}^{\ell+1} - \boldsymbol{u}^\ell}{k} \quad \text{for all } \ell = 0,\dots L-1.$$

For the spatial discretization, we consider a regular tetrahedral triangulation $\mathcal{T}_h$ of Ω with mesh size $h > 0$. We denote by $\mathcal{N}_h$ the set of vertices of $\mathcal{T}_h$ and by $\{\phi_{\boldsymbol{z}}\}_{\boldsymbol{z}\in\mathcal{N}_h}$ the classical nodal basis of the space $\mathcal{S}^1(\mathcal{T}_h)$ of $\mathcal{T}_h$-piecewise linear and globally continuous discrete functions, i.e., $\phi_{\boldsymbol{z}}(\boldsymbol{z}') = \delta_{\boldsymbol{z},\boldsymbol{z}'}$ for all $\boldsymbol{z},\boldsymbol{z}' \in \mathcal{N}_h$. With $\{\boldsymbol{e}_j\}_{j=1,2,3}$ the standard basis of $\mathbb{R}^3$, $\{\phi_{\boldsymbol{z}}\boldsymbol{e}_j\}_{\boldsymbol{z}\in\mathcal{N}_h, j=1,2,3}$ gives a basis of $\mathcal{S}^1(\mathcal{T}_h)^3$. Note that $\mathcal{S}^1(\mathcal{T}_h)^3$ is a $3N$-dimensional space, with N denoting the number of vertices in $\mathcal{N}_h$. We introduce the *set of admissible discrete magnetizations*

$$\mathcal{M}_h := \left\{ \boldsymbol{m}_h \in \mathcal{S}^1(\mathcal{T}_h)^3 : \ |\boldsymbol{m}_h(\boldsymbol{z})| = 1 \text{ for all } \boldsymbol{z} \in \mathcal{N}_h \right\}$$

and, for $\boldsymbol{m}_h \in \mathcal{M}_h$, the *discrete tangent space* of $\boldsymbol{m}_h$

$$\mathcal{K}_h[\boldsymbol{m}_h] := \left\{ \boldsymbol{\varphi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3 : \boldsymbol{m}_h(\boldsymbol{z}) \cdot \boldsymbol{\varphi}_h(\boldsymbol{z}) = 0 \text{ for all } \boldsymbol{z} \in \mathcal{N}_h \right\}.$$

We consider the nodal interpolant $\mathcal{I}_h \colon C^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)$, which is defined by $\mathcal{I}_h(v) = \sum_{\boldsymbol{z}\in\mathcal{N}_h} v(\boldsymbol{z})\phi_{\boldsymbol{z}}$ for all $v \in C^0(\overline{\Omega})$. We denote the vector-valued realization of the nodal interpolant by $\boldsymbol{\mathcal{I}}_h \colon \boldsymbol{C}^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)^3$. In $\boldsymbol{C}^0(\overline{\Omega})$, besides the standard $\boldsymbol{L}^2(\Omega)$-scalar product $\langle\cdot,\cdot\rangle$, we consider the mass-lumped scalar product $\langle\cdot,\cdot\rangle_h$ defined by

$$\langle \boldsymbol{u}, \boldsymbol{w}\rangle_h = \int_\Omega \mathcal{I}_h(\boldsymbol{u}\cdot\boldsymbol{w}) \,\mathrm{d}\boldsymbol{x} \quad \text{for all } \boldsymbol{u},\boldsymbol{w} \in \boldsymbol{C}^0(\overline{\Omega}).$$

Using the definition of the nodal interpolant, we see that

$$\langle \boldsymbol{u}, \boldsymbol{w}\rangle_h = \sum_{\boldsymbol{z}\in\mathcal{N}_h} \beta_{\boldsymbol{z}}\, \boldsymbol{u}(\boldsymbol{z}) \cdot \boldsymbol{w}(\boldsymbol{z}) \quad \text{for all } \boldsymbol{u},\boldsymbol{w} \in \boldsymbol{C}^0(\overline{\Omega}), \tag{10}$$

where $\beta_{\boldsymbol{z}} := \int_\Omega \phi_{\boldsymbol{z}} \,\mathrm{d}\boldsymbol{x} > 0$ for all $\boldsymbol{z} \in \mathcal{N}_h$. For discrete functions, the induced norm $\|\cdot\|_h$ is equivalent to the standard $\boldsymbol{L}^2(\Omega)$-norm; see [9, Lemma 3.9], i.e., it holds that

$$\|\boldsymbol{w}_h\| \le \|\boldsymbol{w}_h\|_h \le \sqrt{5}\,\|\boldsymbol{w}_h\| \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{11}$$

We define the (negative) discrete Laplacian $-\boldsymbol{\Delta}_h \colon \boldsymbol{H}^1(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)^3$ by

$$-\langle \boldsymbol{\Delta}_h \boldsymbol{w}, \boldsymbol{w}_h\rangle_h = \langle \nabla \boldsymbol{w}, \nabla \boldsymbol{w}_h\rangle \quad \text{for all } \boldsymbol{w} \in \boldsymbol{H}^1(\Omega) \text{ and } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{12}$$

Let $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$. With a double application of the classical inverse estimate and the norm equivalence (11), we see that

$$\begin{aligned}\|\boldsymbol{\Delta}_h \boldsymbol{w}_h\|_h^2 &= \langle \boldsymbol{\Delta}_h \boldsymbol{w}_h, \boldsymbol{\Delta}_h \boldsymbol{w}_h\rangle_h \overset{(12)}{=} -\langle \nabla \boldsymbol{w}_h, \nabla \boldsymbol{\Delta}_h \boldsymbol{w}_h\rangle \le \|\nabla \boldsymbol{w}_h\|\, \|\nabla \boldsymbol{\Delta}_h \boldsymbol{w}_h\| \\ &\le C h^{-2}\, \|\boldsymbol{w}_h\|_h\, \|\boldsymbol{\Delta}_h \boldsymbol{w}_h\|_h\,.\end{aligned}$$

This shows that

$$\|\boldsymbol{\Delta}_h \boldsymbol{w}_h\|_h \le C h^{-2}\, \|\boldsymbol{w}_h\|_h \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3, \tag{13}$$

where $C > 0$ depends only on the quasi-uniformity of the triangulation $\mathcal{T}_h$. Finally, we define the mapping $\mathbb{P}_h\colon \boldsymbol{L}^2(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)^3$ by

$$\langle \mathbb{P}_h \boldsymbol{w}, \boldsymbol{w}_h\rangle_h = \langle \boldsymbol{w}, \boldsymbol{w}_h\rangle \quad \text{for all } \boldsymbol{w} \in \boldsymbol{L}^2(\Omega) \text{ and } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{14}$$

Using (10), it is easy to see that, for all $\boldsymbol{w} \in \boldsymbol{L}^2(\Omega)$ and all $\boldsymbol{z} \in \mathcal{N}_h$, it holds that $(\mathbb{P}_h \boldsymbol{w})(\boldsymbol{z}) = \beta_{\boldsymbol{z}}^{-1} \int_\Omega \boldsymbol{w}\phi_{\boldsymbol{z}}\,\mathrm{d}\boldsymbol{x}$. In particular, the computation of $\mathbb{P}_h \boldsymbol{w}$ does not require to solve any linear system.

3. First-order predictor-corrector scheme

In this section, we discuss the first-order scheme proposed in [31] and its connections with the integrators proposed in [10] and [3]. Our contribution is twofold: First, we prove unconditional well-posedness of the scheme, which fills a fundamental gap in the analysis of [31]. Second, we employ an explicit treatment of the (nonlocal) lower-order contributions to obtain a computationally superior IMEX version of the scheme, preserving (unconditional) convergence and experimental rates in time. We first consider the method for the case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m}$. For the general case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$, we refer to Section 3.3.

3.1. Variational formulation

The following algorithm restates [31, Algorithm 1] written in terms of the discrete functions $\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$, where $\boldsymbol{m}_h^\ell \approx \boldsymbol{m}(t_\ell)$, $\boldsymbol{v}_h^\ell \approx \partial_t \boldsymbol{m}(t_\ell)$, and $\boldsymbol{m}_h^{\ell+1} \approx \boldsymbol{m}(t_{\ell+1})$. In particular, the predictor (15) of Algorithm 3.1 reformulates the N equations in $\mathbb{R}^3$ of the predictor of [31, Algorithm 1] as an equivalent variational formulation for $\boldsymbol{v}_h^\ell$ in $\mathcal{S}^1(\mathcal{T}_h)^3$. As for the tangent plane scheme [3], $\theta \in [0,1]$ is a parameter modulating the 'degree of implicitness' of the scheme.

Algorithm 3.1 (*PC1, variational form*). ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: *For all time-steps $\ell = 0, \dots, L-1$, iterate:*

(i) *Compute $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that*

$$\begin{aligned}(1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h = &-\ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell), \boldsymbol{w}_h\rangle_h \\ &- \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell)), \boldsymbol{w}_h\rangle_h\,.\end{aligned} \tag{15}$$

(ii) *Define $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ by*

$$\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})}{\left|\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})\right|} \in \mathbb{S}^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h. \tag{16}$$

Output: *Sequence of discrete functions* $\left\{(\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1})\right\}_{\ell=0}^{L-1}$.

3.2. Unconditional well-posedness

The predictor (15) can be written as: Find $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{w}_h) = F_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}_h) \qquad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3\,,$$

with the linear form $F_{\mathrm{pre}}[\boldsymbol{m}_h^\ell]$ and the bilinear form $a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell]$ on $\mathcal{S}^1(\mathcal{T}_h)^3$ reading

$$\begin{aligned}F_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}_h) &:= -\ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{m}_h^\ell, \boldsymbol{w}_h\rangle_h - \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{m}_h^\ell), \boldsymbol{w}_h\rangle_h\,, \\ a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{w}_h) &:= (1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h + \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h \\ &\quad + \alpha \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{v}_h^\ell), \boldsymbol{w}_h\rangle_h\,.\end{aligned}$$

From the boundedness of $\boldsymbol{m}_h^\ell$ in $\boldsymbol{L}^\infty(\Omega)$ guaranteed by the nodal projection (16) and an inverse estimate on the discrete Laplacian (13) we have

$$a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}_h, \boldsymbol{w}_h) \geq (1 - Ckh^{-2}) \, \|\boldsymbol{w}_h\|_h^2 \,.$$

Hence, assuming the CFL condition $k = o(h^2)$ implies the coercivity of $a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell]$ for sufficiently small h and k. However, this undesirable restriction is a consequence of naively using the inverse estimate, and can be avoided.

For arbitrary $\alpha > 0$ the upcoming refined analysis allows to drop any CFL-type assumptions on the discretization parameters: In Lemma 3.2, we first collect two basic properties of Algorithm 3.1, which turn out to be sufficient to prove unconditional well-posedness of the algorithm in Theorem 3.3; also see Remark 3.4.

Lemma 3.2. *Let $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$. Suppose that the solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ to (15) exists. Then, $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, and (16) provides a well-defined $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.*

Proof. For arbitrary $\boldsymbol{z} \in \mathcal{N}_h$, with $\phi_{\boldsymbol{z}} \in \mathcal{S}^1(\mathcal{T}_h)$ denoting the hat function with $\phi_{\boldsymbol{z}}(\boldsymbol{z}') = \delta_{\boldsymbol{z},\boldsymbol{z}'}$ for all $\boldsymbol{z}' \in \mathcal{N}_h$, we choose $\boldsymbol{w}_h := \boldsymbol{m}_h^\ell(\boldsymbol{z})\phi_{\boldsymbol{z}} \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (15) to see

$$\boldsymbol{m}_h^\ell(\boldsymbol{z}) \cdot \boldsymbol{v}_h^\ell(\boldsymbol{z}) \overset{(10)}{=} \beta_{\boldsymbol{z}}^{-1} \langle \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell(\boldsymbol{z})\phi_{\boldsymbol{z}} \rangle_h \overset{(15),(4b)}{=} 0 \,.$$

Hence, $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ belongs to $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$.
Well-posedness of (16) follows immediately from $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ via

$$|\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})|^2 = |\boldsymbol{m}_h^\ell(\boldsymbol{z})|^2 + k^2|\boldsymbol{v}_h^\ell(\boldsymbol{z})|^2 \geq |\boldsymbol{m}_h^\ell(\boldsymbol{z})|^2 = 1 \quad \text{for all } \ell = 0, \dots, L-1 \,.$$

Consequently, for all $\boldsymbol{z} \in \mathcal{N}_h$ the denominator in (16) is bounded below by $|\boldsymbol{m}_h^\ell(\boldsymbol{z})| = 1$ and the corrector step of Algorithm 3.1 is always well-posed.
The third claim $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ follows directly from the explicit projection in (16). □

These two observations are already sufficient to prove the first main contribution of this work.

Theorem 3.3. *Let $\alpha > 0$. Then, Algorithm 3.1 is unconditionally well-posed for any input $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, i.e., for all $\ell = 0, \dots, L-1$ the predictor (15) admits a unique solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ and the corrector (16) is well-posed providing $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.*

Proof. Well-posedness of the corrector (16) and $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ follow from Lemma 3.2. Transforming (15) into a coercive system in the discrete tangent space, we prove well-posedness of the predictor in three steps:

- **Step 1:** The predictor of Algorithm 3.1 can be reformulated as a well-posed system.

We claim that $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ satisfies (15) for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, if and only if it satisfies $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ as well as

$$\alpha \langle \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h \rangle_h + \langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h \rangle_h = \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h \rangle_h \quad \text{for all } \boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell] \,. \tag{17}$$

This formulation can be written as follows: Find $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ such that

$$a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h) = \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\Delta}_h \boldsymbol{m}_h^\ell, \boldsymbol{\varphi}_h \rangle_h \quad \text{for all } \boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell],$$

where the bilinear form $a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell] \colon \mathcal{K}_h[\boldsymbol{m}_h^\ell] \times \mathcal{K}_h[\boldsymbol{m}_h^\ell] \to \mathbb{R}$ is defined by

$$a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h) := \alpha \langle \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h \rangle_h + \langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h \rangle_h - \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{\Delta}_h \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h \rangle_h.$$

For $\alpha > 0$, the bilinear form satisfies the ellipticity property

$$a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell](\boldsymbol{\varphi}_h, \boldsymbol{\varphi}_h) = \alpha \, \|\boldsymbol{\varphi}_h\|_h^2 + \ell_{\mathrm{ex}}^2 \theta k \, \|\nabla \boldsymbol{\varphi}_h\|^2 \quad \text{for all } \boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell] \,,$$

and the problem (17) is well-posed by the Lax–Milgram theorem. To conclude the proof, it remains to show the claimed equivalence of (15) and (17).

- **Step 2:** Any solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ of (15) also solves (17).

Given arbitrary $\boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, we choose $\boldsymbol{w}_h = \mathcal{I}_h(\alpha\boldsymbol{\varphi}_h + \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell) \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (15) to obtain

$$\begin{aligned}(1+\alpha^2)\alpha\langle \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h + (1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell\rangle_h &= -\alpha\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h\rangle_h \\ &\quad - \ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell\rangle_h - \alpha^2\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell)), \boldsymbol{\varphi}_h\rangle_h \\ &\quad - \alpha\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell)), \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell\rangle_h\,. \end{aligned} \tag{18}$$

By (4d) the left-hand side of (18) resembles the left-hand side of (17) scaled by $(1+\alpha^2)$. From $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, we infer $\mathcal{I}_h(|\boldsymbol{m}_h^\ell|^2) = 1$ and $\mathcal{I}_h(\boldsymbol{m}_h^\ell \cdot \boldsymbol{\varphi}_h) = 0$ in Ω. Hence, using the vector identities (4b)–(4e), the first and the last term on the right-hand side of (18) cancel out, and (18) equivalently reads

$$(1+\alpha^2)\big(\alpha\langle \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h + \langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h\big) = (1+\alpha^2)\ell_{\mathrm{ex}}^2\langle \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h\rangle_h\,.$$

Now multiplying (18) by $1/(1+\alpha^2)$, we conclude that any $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ satisfying (15) necessarily satisfies (17) and, according to Lemma 3.2, belongs to $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$ itself.

- **Step 3:** Any solution $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ of (17) also solves (15).

Given arbitrary $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, we choose $\boldsymbol{\varphi}_h = \mathcal{I}_h\big(\boldsymbol{m}_h^\ell \times \boldsymbol{w}_h + \alpha\boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell)\big) \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ in (17) to obtain

$$\begin{aligned}&\alpha\langle \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times \boldsymbol{w}_h\rangle_h + \alpha^2\langle \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell)\rangle_h \\ &\quad + \langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times \boldsymbol{w}_h\rangle_h + \alpha\langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell)\rangle_h \\ &\quad = \ell_{\mathrm{ex}}^2\langle \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{m}_h^\ell \times \boldsymbol{w}_h\rangle_h + \alpha\ell_{\mathrm{ex}}^2\langle \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell)\rangle_h\,. \end{aligned} \tag{19}$$

From $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, we infer $\mathcal{I}_h(|\boldsymbol{m}_h^\ell|^2) = 1$ and $\mathcal{I}_h(\boldsymbol{m}_h^\ell \cdot \boldsymbol{v}_h^\ell) = 0$ in Ω. Hence, by the vector identities (4b)–(4e), the first and the last term on the left-hand side of (19) cancel out, while the second and third term on the left-hand side of (19) add up to the left-hand side of (15). Further, by (4d) the right-hand side of (19) resembles the right-hand side of (15). We conclude that any $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell] \subset \mathcal{S}^1(\mathcal{T}_h)^3$ satisfying (17) necessarily satisfies (15). Ultimately, we have shown that (15) is equivalent to (17), which always allows for a unique solution as shown in Step 1. □

Remark 3.4. (i) Let $\boldsymbol{w}\colon \Omega \to \mathbb{R}^3$ be an arbitrary smooth test function. Writing $\boldsymbol{m}^\ell := \boldsymbol{m}(t_\ell)$ and $\boldsymbol{v}^\ell := \partial_t\boldsymbol{m}(t_\ell)$, the variational formulation of the LL form (5a) of LLG at time $t_\ell \in (0,T)$ reads

$$(1+\alpha^2)\langle \boldsymbol{v}^\ell, \boldsymbol{w}\rangle \;=\; -\ell_{\mathrm{ex}}^2\,\langle \boldsymbol{m}^\ell \times \boldsymbol{\Delta}\boldsymbol{m}^\ell, \boldsymbol{w}\rangle - \alpha\ell_{\mathrm{ex}}^2\,\langle \boldsymbol{m}^\ell \times (\boldsymbol{m}^\ell \times \boldsymbol{\Delta}\boldsymbol{m}^\ell), \boldsymbol{w}\rangle\,. \tag{20}$$

The discrete variational formulation (15) can be seen as a discrete mass-lumped version of (20), where the effective field is treated implicitly in time.
(ii) The core of the proof of Theorem 3.3 is the equivalent reformulation of the predictor step (15) as well-posed system (17) in the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$. For $\alpha > 0$, the reformulated system is unconditionally well-posed and corresponds to a discretization of the alternative form of LLG (3). Using (4c) and $|\boldsymbol{m}|^2 \equiv 1$, the formulation (3) is directly obtained from the LL form (5a) via $(\alpha\cdot(5a) + \boldsymbol{m}\times(5a))/(1+\alpha^2)$. Step 3 of the proof of Theorem 3.3 resembles the analogous computations on a discrete level. We emphasize, that the mass-lumped scalar product $\langle\cdot,\cdot\rangle_h$ as well as $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ are the crucial ingredients in the proof of Theorem 3.3.
(iii) With the reformulation (17), we fully understand the real nature of the first-order integrator from [31]: It is a predictor-corrector scheme which combines the approaches of Bartels & Prohl [10] (mass-lumping (10), discrete Laplacian (12)) and Alouges [3] (degree of implicitness θ, projection update (16), unknown approximates time derivative). The predictor step (15) is a mass-lumped discrete variational formulation of the LL form (5a) of LLG. The equivalent variational formulation (17) is a mass-lumped variational formulation of the alternative form (3) of LLG and, in particular, is the mass-lumped version of the predictor step of the tangent plane scheme from [3]. Analogously to the tangent plane scheme, the corrector step of Algorithm 3.1 employs the nodal projection to enforce the modulus constraint at the vertices of the triangulations.
(iv) While the proof of Theorem 3.3 emphasizes the close relation of Algorithm 3.1 to the first-order tangent plane scheme, it is restricted to $\alpha > 0$. In fact, Theorem 3.3 can also be proved for the limit case $\alpha = 0$; see Remark 4.6(iii)–(iv) below.

3.3. Including lower-order contributions

In this section, we discuss the extension of the scheme to the general case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2\,\boldsymbol{\Delta}\boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$. We start by recalling the definition (14) of the mapping $\mathbb{P}_h\colon \boldsymbol{L}^2(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)^3$ and assume that we are given an operator $\boldsymbol{\pi}_h\colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ which approximates $\boldsymbol{\pi}$, e.g., in the case of the nonlocal stray field $\boldsymbol{\pi}(\boldsymbol{m}) = \boldsymbol{h}_{\mathrm{s}}$, $\boldsymbol{\pi}_h$ is a method for the approximation of the magnetostatic Maxwell equations, e.g., via the hybrid finite element / boundary element method (FEM-BEM) from [25]. This approach exploits the special structure of the full-space equation and requires only the sequential solution of two FEM

problems for the interior Poisson equation, where inhomogeneous Dirichlet conditions are obtained with a double-layer boundary integral operator. We refer to [38, Algorithm 12] for a precise statement of the hybrid FEM-BEM method directly applicable to our setting (and used in the numerical experiments in Section 5 below).

In the original first-order integrator from [31], the lower-order contributions are treated implicitly in time. Rewritten as a mass-lumped discrete LL formulation like (15), the predictor step of [31, Algorithm 1] reads as follows: Find $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$\begin{aligned}(1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h &= -\langle \boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+\theta})], \boldsymbol{w}_h\rangle_h \\ &\quad - \alpha \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+\theta})]), \boldsymbol{w}_h\rangle_h \end{aligned} \tag{21}$$

for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$. Here, $\boldsymbol{f}^{\ell+\theta} = \boldsymbol{f}(t_\ell + \theta k)$ for all $\ell = 0, \dots, L-1$. However, this approach for the inclusion of the lower-order terms is not very attractive from the computational point of view: Indeed, the variational formulation comprises the term $\boldsymbol{\pi}_h(\boldsymbol{v}_h^\ell)$ which requires to solve a (possibly nonlocal) problem for the unknown. An implementation of this scheme would then be based on a costly inner iteration.

From our previous work on the tangent plane scheme [14,21] and on the midpoint scheme [38], we know that an explicit treatment is favorable: Therefore, we change the above variational formulation: Find $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$\begin{aligned}(1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h &= -\langle \boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^{\ell})], \boldsymbol{w}_h\rangle_h \\ &\quad - \alpha \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^{\ell})]), \boldsymbol{w}_h\rangle_h \end{aligned}$$

for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$. Only the leading-order exchange contribution is treated implicitly in time, while the lower-order contributions are treated explicitly. This does not spoil the convergence result of the scheme (since the nodal projection already restricts the scheme to first-order in time) and it is computationally much more attractive. To sum up, we consider the following implicit-explicit (IMEX) algorithm.

Algorithm 3.5 (*PC1+IMEX*). ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: *For all time-steps* $\ell = 0, \dots, L-1$, *iterate:*

(i) *Compute* $\mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell)) \in \mathcal{S}^1(\mathcal{T}_h)^3$.
(ii) *Compute* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$\begin{aligned}(1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h &= -\langle \boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^{\ell})], \boldsymbol{w}_h\rangle_h \\ &\quad - \alpha \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^{\ell})]), \boldsymbol{w}_h\rangle_h . \end{aligned} \tag{22}$$

(iii) *Define* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ *by*

$$\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})}{|\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h. \tag{23}$$

Output: *Sequence of discrete functions* $\big\{(\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1})\big\}_{\ell=0}^{L-1}$.

3.4. Stability of Algorithm 3.5

Well-posedness of Algorithm 3.5 follows from well-posedness of Algorithm 3.1 (Theorem 3.3), as the system matrices for the linear systems corresponding to the left-hand sides of (22) and (15), respectively, coincide.

For stability of Algorithm 3.5, we assume that all off-diagonal entries of the stiffness matrix $A = (a_{\boldsymbol{z},\boldsymbol{z}'})_{\boldsymbol{z},\boldsymbol{z}' \in \mathcal{N}_h}$ are non-positive, i.e., it holds that

$$a_{\boldsymbol{z},\boldsymbol{z}'} = \langle \nabla\phi_{\boldsymbol{z}'}, \nabla\phi_{\boldsymbol{z}}\rangle \le 0 \quad \text{for all } \boldsymbol{z}, \boldsymbol{z}' \in \mathcal{N}_h \text{ with } \boldsymbol{z} \ne \boldsymbol{z}'. \tag{24}$$

This requirement, usually referred to as *angle condition*[1], ensures that the nodal projection $\boldsymbol{w}_h \mapsto \mathcal{I}_h\big[\boldsymbol{w}_h/|\boldsymbol{w}_h|\big]$ does not increase the exchange energy of a discrete function, i.e., it holds that

$$\big\|\nabla \mathcal{I}_h\big[\boldsymbol{w}_h/|\boldsymbol{w}_h|\big]\big\| \le \|\nabla \boldsymbol{w}_h\| , \tag{25}$$

for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$ with $|\boldsymbol{w}_h(\boldsymbol{z})| \ge 1$ for all $\boldsymbol{z} \in \mathcal{N}_h$; see [8, Lemma 3.2]. Moreover, we assume that the discrete operator $\boldsymbol{\pi}_h\colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \boldsymbol{L}^2(\Omega)$ is stable in the sense that

[1] The assumption (24) is usually referred to as angle condition, because in 3D it is satisfied, e.g., if all dihedral angles of all tetrahedra of $\mathcal{T}_h$ are $\le \pi/2$.

$$\|\boldsymbol{\pi}_h(\boldsymbol{w}_h)\| \le C\,\|\boldsymbol{w}_h\| \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3\,, \tag{26}$$

which is met in many practical situations; see [14]. Under these assumptions, there holds stability of Algorithm 3.5.

Theorem 3.6. *Let $\mathcal{T}_h$ such that (25) holds true. For input $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, let $\big\{(\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1})\big\}_{\ell=0}^{L-1}$ be the output of Algorithm 3.5. Then, for all $J = 0, \dots, L-1$, there holds the stability estimate*

$$\begin{aligned}&\frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^J\right\|^2 + \alpha k \sum_{\ell=0}^{J-1}\left\|\boldsymbol{v}_h^\ell\right\|^2 + \ell_{\mathrm{ex}}^2(\theta-1/2)k^2\sum_{\ell=0}^{J-1}\left\|\nabla \boldsymbol{v}_h^\ell\right\|^2\\ &\quad\le \frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^0\right\|^2 + k\sum_{\ell=0}^{J-1}\langle \boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell\rangle.\end{aligned} \tag{27}$$

Proof. To abbreviate notation we define

$$\boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell) := \ell_{\mathrm{ex}}^2\boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell) \in \mathcal{S}^1(\mathcal{T}_h)^3\,.$$

Testing (22) with $\boldsymbol{w}_h = \boldsymbol{v}_h^\ell$, $\boldsymbol{w}_h = \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)$, and $\boldsymbol{w}_h = \mathcal{I}_h(\boldsymbol{m}_h^\ell \times \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell))$, respectively, leads to

$$(1+\alpha^2)\left\|\boldsymbol{v}_h^\ell\right\|_h^2 = \langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h + \alpha\langle \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h, \tag{28a}$$

$$\alpha\left\|\boldsymbol{m}_h^\ell \times \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\right\|_h^2 = (1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h\,, \tag{28b}$$

$$\left\|\boldsymbol{m}_h^\ell \times \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\right\|_h^2 = (1+\alpha^2)\langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h\,, \tag{28c}$$

where we used $\mathcal{I}_h(|\boldsymbol{m}_h^\ell|^2) = 1$ and $\mathcal{I}_h(\boldsymbol{m}_h^\ell \cdot \boldsymbol{v}_h^\ell) = 0$ in Ω together with the identities (4b)–(4e). Combining (28a)–(28c) gives

$$\alpha\left\|\boldsymbol{v}_h^\ell\right\|_h^2 = \langle \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h\,.$$

Plugging in the definition of $\boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)$, we see

$$\ell_{\mathrm{ex}}^2\langle \nabla \boldsymbol{v}_h^\ell, \nabla \boldsymbol{m}_h^\ell\rangle = -\alpha\left\|\boldsymbol{v}_h^\ell\right\|_h^2 - \ell_{\mathrm{ex}}^2\theta k\left\|\nabla \boldsymbol{v}_h^\ell\right\|^2 + \langle \boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell\rangle. \tag{29}$$

Using the angle condition, we deduce that

$$\begin{aligned}&\frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^{\ell+1}\right\|^2 - \frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^\ell\right\|^2 \overset{(25)}{\le} \frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla(\boldsymbol{m}_h^\ell + k\boldsymbol{v}_h^\ell)\right\|^2 - \frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^\ell\right\|^2\\ &\quad= \ell_{\mathrm{ex}}^2 k\langle \nabla \boldsymbol{m}_h^\ell, \nabla \boldsymbol{v}_h^\ell\rangle + \frac{\ell_{\mathrm{ex}}^2}{2}k^2\left\|\nabla \boldsymbol{v}_h^\ell\right\|^2\\ &\quad\overset{(29)}{=} -\alpha k\left\|\boldsymbol{v}_h^\ell\right\|_h^2 - \ell_{\mathrm{ex}}^2(\theta-1/2)k^2\left\|\nabla \boldsymbol{v}_h^\ell\right\|^2 + k\langle \boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell\rangle.\end{aligned}$$

Summing over $\ell = 0, \dots, J-1$, we obtain that

$$\begin{aligned}&\frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^J\right\|^2 + \alpha k\sum_{\ell=0}^{J-1}\left\|\boldsymbol{v}_h^\ell\right\|_h^2 + \ell_{\mathrm{ex}}^2(\theta-1/2)k^2\sum_{\ell=0}^{J-1}\left\|\nabla \boldsymbol{v}_h^\ell\right\|^2\\ &\quad\le \frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^0\right\|^2 + k\sum_{\ell=0}^{J-1}\langle \boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell\rangle.\end{aligned}$$

Finally, the norm equivalence (11) yields (27). □

Remark 3.7. (i) The stability (27) is the very same estimate that one obtains for the first-order tangent plane scheme from [3]; see, e.g., [14, Lemma 3.5]. Combining this estimate with the stability of $\boldsymbol{\pi}_h$ from (26), one obtains boundedness of the discrete solutions, which allows to apply the standard compactness argument for parabolic PDEs to prove convergence; see, e.g., [3, Section 3] or [14, Section 3.5].

(ii) Consequently, for both Algorithm 3.1 and Algorithm 3.5, one obtains a convergence result identical to [3, Theorem 2, Remark 1]. In particular, as $h, k \to 0$, for $1/2 < \theta \leq 1$ no coupling of the discretization parameters is necessary, while the CFL conditions $k = o(h)$ and $k = o(h^2)$ are proved to be sufficient for $\theta = 1/2$ and $0 \leq \theta < 1/2$, respectively.
(iii) We note that [31, Theorem 2.2] and its proof are slightly inaccurate and, in particular, the CFL condition $k = o(h)$ is missing for $\theta = 1/2$.

We briefly comment on a projection-free modification of `PC1+IMEX`.

Remark 3.8. As pointed out in Remark 3.4, Algorithm 3.5 and the first-order tangent-plane scheme from [3] coincide up to mass-lumped integration in the predictor (22). Hence, an obvious modification of Algorithm 3.5 in the spirit of the projection-free tangent-plane scheme from [1, Algorithm 6] is omitting the projection in the corrector (23), i.e., defining $\boldsymbol{m}_h^{\ell+1} := \boldsymbol{m}_h^\ell + k\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$. For this projection-free variant of Algorithm 3.5, at first glance, one could hope for the same desirable theoretical features as for the projection-free tangent plane scheme — namely stability and weak convergence [1] without the angle condition (24) and even strong convergence [23], both at the price of a slight deterioration from nodewise unit-length $\boldsymbol{m}_h^\ell \notin \mathcal{M}_h$. In contrast to the projection-free tangent plane scheme, the projection-free variant of Algorithm 3.5 is unconditionally well-posed even for the limit case $\alpha = 0$; see Remark 4.6(iii)–(iv) below. Further, it satisfies a discrete energy law, which, e.g., in the exchange-only case for $\theta = 1/2$ reads

$$\frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^J\right\|^2 + \frac{\alpha}{1+\alpha^2}\ell_{\mathrm{ex}}^4 k \sum_{\ell=0}^{J-1}\left\|\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)\right\|_h^2 = \frac{\ell_{\mathrm{ex}}^2}{2}\left\|\nabla \boldsymbol{m}_h^0\right\|^2 .$$

However, due to the loss of nodewise unit-length $\boldsymbol{m}_h^\ell \notin \mathcal{M}_h$, equivalence of the predictor of the projection-free version of Algorithm 3.5 and the discrete tangent space system (17) in $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$ does not hold anymore. Consequently, the analysis for the projection-free tangent plane scheme from [1,23] does not (directly) transfer, and a rigorous analysis of the projection-free version of Algorithm 3.5 remains open.

4. Second-order predictor-corrector scheme

In this section, we discuss the second-order scheme proposed in [31]. Our contribution is threefold: In theory, well-posedness (for the predictor) of the scheme (which was left open in [31]) follows already from our analysis in Section 3.2. When accounting for the use of inexact (iterative) linear solvers, which is inevitable in practice, however, discrete unit-length $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ is lost and therefore a conceptually new analysis is required to guarantee well-posedness in practice. We fill this fundamental gap in the analysis of [31] for their second-order scheme, by proving unconditional well-posedness not only for the proposed predictor-corrector scheme, but also for its practical version incorporating inexact (iterative) linear solvers. Again, we first consider the method for the case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m}$. The general case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$ is treated in Section 4.4, where we employ an explicit treatment of the (nonlocal) lower-order contributions to obtain a computationally superior IMEX version of the scheme, preserving experimental rates in time. We numerically confirm the applicability and the formal second-order of the proposed IMEX scheme in Section 5. Theoretical stability (and hence convergence) of the second-order scheme remains open (like in [31]), but is experimentally investigated in a numerical study in Section 5.3.

4.1. Variational formulation

The following algorithm restates [31, Algorithm 2] written in terms of the discrete functions $\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$. In particular, the corrector (31) of Algorithm 4.1 reformulates the N equations in $\mathbb{R}^3$ of the corrector of [31, Algorithm 2] as an equivalent variational formulation for $\boldsymbol{m}_h^{\ell+1}$ in $\mathcal{S}^1(\mathcal{T}_h)^3$. The predictor step coincides with step (i) of Algorithm 3.1, i.e., (30) coincides with (15). As in Section 3, the parameter $\theta \in [0,1]$ modulates the 'degree of implicitness' (in the predictor) of the scheme.

Algorithm 4.1 (*`PC2`, variational form*). ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: *For all time-steps* $\ell = 0, \dots, L-1$, *iterate:*

(i) *Compute* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$\begin{aligned}(1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h = &-\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell), \boldsymbol{w}_h\rangle_h \\ &- \alpha\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell)), \boldsymbol{w}_h\rangle_h .\end{aligned} \tag{30}$$

(ii) *Compute* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$(1+\alpha^2)\langle d_t \boldsymbol{m}_h^{\ell+1}, \boldsymbol{w}_h\rangle_h = -\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^{\ell+1/2} \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell), \boldsymbol{w}_h\rangle_h \tag{31}$$

$$- \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^{\ell+1/2} \times [(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)], \boldsymbol{w}_h \rangle_h \,.$$

Output: *Sequence of discrete functions* $\big\{(\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1})\big\}_{\ell=0}^{L-1}$.

The corrector step of Algorithm 3.1, which combines a linear first-order time-stepping with the nodal projection, is replaced by the linear system (31). The 2D numerical results of [31, Figure 3] indicate that the method is of second-order in time. In Section 5.2, we confirm this observation for a numerical example in 3D.

4.2. Unconditional well-posedness, exact solver

In Lemma 4.2, we first collect two basic properties of Algorithm 4.1, which, for $\alpha > 0$, turn out to be sufficient to prove unconditional well-posedness of the algorithm in Theorem 4.3.

Lemma 4.2. *Let* $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$. *Suppose that the solutions* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ *and* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$ *to* (30) *and* (31) *exist, respectively. Then,* $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, *and* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.

Proof. The claim $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ follows as in the proof of Lemma 3.2. We show that $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ implies $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ due to the corrector system (31): For arbitrary $\boldsymbol{z} \in \mathcal{N}_h$, we choose $\boldsymbol{w}_h := \boldsymbol{m}_h^{\ell+1/2}(\boldsymbol{z})\phi_{\boldsymbol{z}} \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (31) to see

$$\frac{(1+\alpha^2)\beta_{\boldsymbol{z}}}{2k}\Big(|\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z})|^2 - |\boldsymbol{m}_h^\ell(\boldsymbol{z})|^2\Big) \overset{(10)}{=} (1+\alpha^2)\langle d_t \boldsymbol{m}_h^{\ell+1}, \boldsymbol{m}_h^{\ell+1/2}(\boldsymbol{z})\phi_{\boldsymbol{z}}\rangle_h \overset{(31),(4b)}{=} 0\,.$$

This shows that $|\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z})| = |\boldsymbol{m}_h^\ell(\boldsymbol{z})|$ for all $\boldsymbol{z} \in \mathcal{N}_h$. Hence, $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ implies that $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$. The assumption $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ concludes the proof. □

We show unconditional well-posedness of the corrector (31), while with Lemma 4.2 unconditional well-posedness of the predictor is inferred from our analysis in Section 3.2.

Theorem 4.3. *Let* $\alpha > 0$. *Then, Algorithm 4.1 is unconditionally well-posed for any input* $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, *i.e., for all* $\ell = 0, \dots, L-1$, *the predictor* (30) *admits a unique solution* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$, *and the corrector* (31) *admits a unique solution* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.

Proof. By Lemma 4.2 it holds that $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ for all $\ell = 0, \dots, L-1$. Hence, as for the predictor of Algorithm 3.1, the predictor system (30) is equivalent to a coercive system in the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$ with unique solution $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$; see (the proof of) Theorem 3.3. It remains to show well-posedness of the corrector (31): We rewrite the problem in terms of the unknown $\boldsymbol{\eta}_h^\ell := \boldsymbol{m}_h^{\ell+1/2}$, which, by construction, satisfies that $\boldsymbol{m}_h^{\ell+1} = 2\boldsymbol{\eta}_h^\ell - \boldsymbol{m}_h^\ell$ and $d_t \boldsymbol{m}_h^{\ell+1} = 2(\boldsymbol{\eta}_h^\ell - \boldsymbol{m}_h^\ell)/k$. The corrector system (31) then reads: Find $\boldsymbol{\eta}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell](\boldsymbol{\eta}_h^\ell, \boldsymbol{w}_h) = (1+\alpha^2)\langle \boldsymbol{m}_h^\ell, \boldsymbol{w}_h\rangle_h,$$

where the bilinear form $a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell]\colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$ is defined by

$$\begin{aligned} a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell](\boldsymbol{\eta}_h^\ell, \boldsymbol{w}_h) &:= (1+\alpha^2)\langle \boldsymbol{\eta}_h^\ell, \boldsymbol{w}_h\rangle_h + \frac{\ell_{\mathrm{ex}}^2 k}{2}\langle \boldsymbol{\eta}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell), \boldsymbol{w}_h\rangle_h \\ &\quad + \frac{\alpha \ell_{\mathrm{ex}}^2 k}{2}\langle \boldsymbol{\eta}_h^\ell \times [(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)], \boldsymbol{w}_h\rangle_h. \end{aligned}$$

As the bilinear form satisfies the ellipticity property

$$a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell](\boldsymbol{w}_h, \boldsymbol{w}_h) = (1+\alpha^2)\,\|\boldsymbol{w}_h\|_h^2 \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3,$$

the problem is well-posed by the Lax–Milgram theorem. Hence, (31) provides a unique solution $\boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$. Lemma 4.2 guarantees $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ concluding the proof. □

Remark 4.4. (i) Algorithm 4.1 is a predictor-corrector scheme: Both systems, for the predictor (30) and for the corrector (31), respectively, are linear systems representing discrete mass-lumped variational versions of the LL form (5a) of LLG; see also Remark 3.4(i). First, treating the effective field implicitly in time, an approximate time derivative $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, the predictor, is computed. In the second step (the effective field of) the predicted midpoint $\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ is used to compute a corrected update $d_t \boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$, guaranteeing conservation of discrete unit-length $\boldsymbol{m}_h^{\ell+1} := \boldsymbol{m}_h^\ell + k d_t \boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.
(ii) In the proof of Theorem 4.3, note that the assumption $\alpha > 0$ is only exploited to apply Theorem 3.3. Hence, analogously to Theorem 3.3 (Remark 3.4(iv)), also Theorem 4.3 can be extended to the limit case $\alpha = 0$; see Theorem 4.5 below.

4.3. Unconditional well-posedness, inexact solver

Considering the effect of numerical approximations, we extend the theoretical well-posedness result from the previous section to the practical case.

Well-posedness of the predictor step (i) of Algorithm 4.1 is guaranteed by Theorem 3.3: There, under the crucial condition $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$, computing $\boldsymbol{v}_h^\ell$ in the predictor step is shown to be equivalent to solving the system (17) in the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$, which is always well-posed for $\alpha > 0$. While $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ is explicitly enforced in step (ii) of Algorithm 3.1, in Algorithm 4.1 it follows only implicitly from the inherent length preservation guaranteed by the variational formulation (31) solved in step (ii) together with $\boldsymbol{m}_h^{\ell-1} \in \mathcal{M}_h$ in the previous time-step; see the proof of Lemma 4.2. In practice however, linear systems are solved by inexact (iterative) numerical solvers, i.e., the coefficient vector of the unknown $\boldsymbol{m}_h^{\ell+1}$ solves the linear system of equations corresponding to (31) only up to some accuracy $\varepsilon > 0$, commonly in the $\ell^2(\mathbb{R}^{3N})$-norm. Consequently, for any $\boldsymbol{z} \in \mathcal{N}_h$ there only holds $|\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z})| \approx |\boldsymbol{m}_h^\ell(\boldsymbol{z})|$ with a small error depending on the discretization parameters ε and h. Moreover, the deviation from nodewise unit-length accumulates over the time-steps $\ell = 0, \dots, L-1$. Consequently — if recoverable at all — one expects to require CFL-type couplings of the discretization parameters k, h, ε to rigorously argue (approximate) equivalence of the linear system in step (i) of Algorithm 4.1 and the well-posed system (17) in the proof of Theorem 3.3.

To avoid these analytical difficulties, we take a different analytical approach: The new analysis uses a space decomposition technique reformulating (30) as an equivalent saddle-point problem, which subsequently is proved to be unconditionally well-posed and hence always provides a unique solution. In particular, this does not require $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$, but allows for arbitrary $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3 \supsetneqq \mathcal{M}_h$. Additionally, the analysis applies to all $\alpha \geq 0$, extending well-posedness of Algorithm 4.1 to the Schrödinger map equation ($\alpha = 0$).

Theorem 4.5. *Let $\alpha \geq 0$. Then, Algorithm 4.1 is unconditionally well-posed for any input $\boldsymbol{m}_h^0 \in \mathcal{S}^1(\mathcal{T}_h)^3$, i.e., for all $\ell = 0, \dots, L-1$ and any $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$, the predictor (30) admits a unique solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$, and the corrector (31) admits a unique solution $\boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$.*

Proof. For arbitrary $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ well-posedness of the corrector (31) is guaranteed by the proof of Theorem 4.3, as it does not require $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$. Using a space decomposition technique, we show unconditional well-posedness of the predictor system (30) for any $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ — in particular for $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ not necessarily belonging to $\mathcal{M}_h$ — in five steps:

- **Step 0:** Some notation.

Throughout, for an operator $A\colon X \to Y$ between two Hilbert spaces, we write $\mathcal{R}(A) \subseteq Y$ for its range, and $\mathcal{N}(A) \subseteq X$ for its kernel. We consider the (negative) discrete Laplace operator (12) restricted to $\mathcal{S}^1(\mathcal{T}_h)^3 \subset \boldsymbol{H}^1(\Omega)$, which will be denoted by the same symbol $-\boldsymbol{\Delta}_h\colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathcal{S}^1(\mathcal{T}_h)^3$. Further, we identify a 3-vector with the corresponding constant vector-valued grid function, i.e., $\mathbb{R}^3 \subset \big(\mathcal{S}^1(\mathcal{T}_h)^3, \langle\cdot,\cdot\rangle_h\big)$. For $S \subset \mathcal{S}^1(\mathcal{T}_h)^3$ a subspace we denote by $\boldsymbol{I}_S$ the identity on S.

- **Step 1:** Orthogonal decomposition $\mathcal{S}^1(\mathcal{T}_h)^3 = \mathcal{R}(\boldsymbol{P}_*) \oplus \mathcal{N}(\boldsymbol{P}_*)$.

Define the operator $\boldsymbol{P}_*\colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathcal{S}^1(\mathcal{T}_h)^3$ for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$ via

$$(\boldsymbol{P}_*\boldsymbol{w}_h)_j = (\boldsymbol{w}_h)_j - \operatorname{meas}(\Omega)^{-1}\langle \boldsymbol{w}_h, \boldsymbol{e}_j\rangle_h \in \mathcal{S}^1(\mathcal{T}_h) \qquad \text{for all} \quad j = 1,2,3\,.$$

Clearly, $\boldsymbol{P}_*$ is the $\langle\cdot,\cdot\rangle_h$-orthogonal projector onto

$$\mathcal{R}(\boldsymbol{P}_*) = \mathcal{S}^1_*(\mathcal{T}_h)^3 := \{\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3\colon \langle \boldsymbol{w}_h, \boldsymbol{e}_j\rangle_h = 0 \text{ for all } j = 1,2,3\}\,,$$

the subset of $\mathcal{S}^1(\mathcal{T}_h)^3$ consisting of the vector-valued grid functions which have zero mean in each component. Due to self-adjointness, $\boldsymbol{P}_*$ provides the orthogonal decomposition

$$\mathcal{S}^1(\mathcal{T}_h)^3 = \mathcal{R}(\boldsymbol{P}_*) \oplus \mathcal{N}(\boldsymbol{P}_*) = \mathcal{S}^1_*(\mathcal{T}_h)^3 \oplus \mathbb{R}^3\,.$$

With respect to this decomposition, we rewrite the unknown $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ as the orthogonal sum

$$\boldsymbol{v}_h^\ell = \boldsymbol{P}_*\boldsymbol{v}_h^\ell \oplus (\boldsymbol{I}_{\mathcal{S}^1(\mathcal{T}_h)^3} - \boldsymbol{P}_*)\boldsymbol{v}_h^\ell =: \boldsymbol{v}_* \oplus \overline{\boldsymbol{v}}\,, \tag{32}$$

with unique $\boldsymbol{v}_* \in \mathcal{R}(\boldsymbol{P}_*) = \mathcal{S}^1_*(\mathcal{T}_h)^3$ and $\overline{\boldsymbol{v}} \in \mathcal{N}(\boldsymbol{P}_*) = \mathbb{R}^3$. Note, that $\overline{\boldsymbol{v}} \in \mathbb{R}^3$ is the vector-valued mean of $\boldsymbol{v}_h^\ell$, i.e., $\langle \overline{\boldsymbol{v}}, \boldsymbol{e}_j\rangle_h = \langle \boldsymbol{v}_h^\ell, \boldsymbol{e}_j\rangle_h$ for all components $j = 1,2,3$.

- **Step 2:** Reduced operator $-\widetilde{\boldsymbol{\Delta}}_h\colon \mathcal{S}^1_*(\mathcal{T}_h)^3 \to \mathcal{S}^1_*(\mathcal{T}_h)^3$.

The discrete Laplacian $-\boldsymbol{\Delta}_h\colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathcal{S}^1(\mathcal{T}_h)^3$ is linear, self-adjoint and by definition (12) has the kernel $\mathcal{N}(-\boldsymbol{\Delta}_h) = \mathbb{R}^3 \subset \mathcal{S}^1(\mathcal{T}_h)^3$. Hence, there holds the orthogonal decomposition

$$\mathcal{S}^1(\mathcal{T}_h)^3 = \mathcal{R}(-\boldsymbol{\Delta}_h) \oplus \mathcal{N}(-\boldsymbol{\Delta}_h) = \mathcal{N}(-\boldsymbol{\Delta}_h)^\perp \oplus \mathcal{N}(-\boldsymbol{\Delta}_h) = \mathcal{S}^1_*(\mathcal{T}_h)^3 \oplus \mathbb{R}^3\,. \tag{33}$$

Consequently, the reduced operator $-\boldsymbol{\Delta}_h|_{\mathcal{S}^1_*(\mathcal{T}_h)^3} =: -\widetilde{\boldsymbol{\Delta}}_h\colon \mathcal{S}^1_*(\mathcal{T}_h)^3 \to \mathcal{S}^1_*(\mathcal{T}_h)^3$ is linear, self-adjoint, and bijective. Moreover, it provides a well-defined inverse denoted by $(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\colon \mathcal{S}^1_*(\mathcal{T}_h)^3 \to \mathcal{S}^1_*(\mathcal{T}_h)^3$ with the same attributes. We point out the identities

$$(-\widetilde{\boldsymbol{\Delta}}_h)^{-1} \circ (-\boldsymbol{\Delta}_h) = \boldsymbol{P}_* \quad \text{and} \quad -\boldsymbol{\Delta}_h \circ (-\widetilde{\boldsymbol{\Delta}}_h)^{-1} = \boldsymbol{P}_*|_{\mathcal{S}^1_*(\mathcal{T}_h)^3} = \boldsymbol{I}_{\mathcal{S}^1_*(\mathcal{T}_h)^3}\,, \tag{34}$$

which follow from the orthogonal decomposition (33).

- **Step 3:** Equivalent saddle point formulation.

With the unknowns $\boldsymbol{q} := -\boldsymbol{\Delta}_h \boldsymbol{v}_* \in \mathcal{S}^1_*(\mathcal{T}_h)^3$ and $\boldsymbol{\lambda} := \overline{\boldsymbol{v}} \in \mathbb{R}^3$ from (32), we induce the representation $\boldsymbol{v}_h^\ell = (-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q} \oplus \boldsymbol{\lambda}$. Plugging this identity into (30), we rewrite the predictor as equivalent saddle point problem: Find $(\boldsymbol{q}, \boldsymbol{\lambda}) \in \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3$, such that for all $(\boldsymbol{w}, \boldsymbol{\mu}) \in \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3$ it holds that

$$a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{q}, \boldsymbol{w}) + b_{\mathrm{sp}}(\boldsymbol{w}, \boldsymbol{\lambda}) = F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{w})\,, \tag{35a}$$

$$b_{\mathrm{sp}}(\boldsymbol{q}, \boldsymbol{\mu}) = 0\,, \tag{35b}$$

with the (bi-)linear forms $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]\colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$, $b\colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3 \to \mathbb{R}$, and $F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]\colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$ given by

$$\begin{aligned}
a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{q}, \boldsymbol{w}) &:= (1+\alpha^2)\langle (-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{P}_*\boldsymbol{q}, \boldsymbol{w}\rangle_h \\
&\quad - \ell_{\mathrm{ex}}^2\theta k\langle \boldsymbol{m}_h^\ell \times \boldsymbol{q}, \boldsymbol{w}\rangle_h - \alpha\ell_{\mathrm{ex}}^2\theta k\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{q}), \boldsymbol{w}\rangle_h\,, \\
b_{\mathrm{sp}}(\boldsymbol{w}, \boldsymbol{\lambda}) &:= (1+\alpha^2)\langle \boldsymbol{\lambda}, \boldsymbol{w}\rangle_h\,, \\
F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}) &:= -\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h\boldsymbol{m}_h^\ell, \boldsymbol{w}\rangle_h - \alpha\ell_{\mathrm{ex}}^2\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h\boldsymbol{m}_h^\ell), \boldsymbol{w}\rangle_h\,.
\end{aligned}$$

The equivalence of (35a)–(35b) to (30) follows from $\boldsymbol{\lambda} \in \mathcal{N}(-\boldsymbol{\Delta}_h)$ and (34). We use the operator $(-\widetilde{\boldsymbol{\Delta}}_h)^{-1} \circ \boldsymbol{P}_*$ rather than $(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}$, so that the bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ is well-defined on $\mathcal{S}^1(\mathcal{T}_h)^3 \supsetneq \mathcal{S}^1_*(\mathcal{T}_h)^3$. The second equation (35b) ensures $\boldsymbol{q} \in \mathcal{S}^1_*(\mathcal{T}_h)^3$, which is not enforced explicitly.

- **Step 4:** The bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ is coercive on the kernel of b_{sp}.

We aim to apply the Brezzi theory for saddle point problems; see, e.g., [12, Section 4.2]. Hence, we require coercivity of the bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]\colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$ on

$$\begin{aligned}
\bigcap_{\boldsymbol{\lambda}\in\mathbb{R}^3} \mathcal{N}\big(b_{\mathrm{sp}}(\cdot, \boldsymbol{\lambda})\big) &= \bigcap_{\boldsymbol{\lambda}\in\mathbb{R}^3} \{\boldsymbol{w} \in \mathcal{S}^1(\mathcal{T}_h)^3\colon \langle \boldsymbol{\lambda}, \boldsymbol{w}\rangle_h = 0\} \\
&= \bigcap_{j=1,2,3} \{\boldsymbol{w} \in \mathcal{S}^1(\mathcal{T}_h)^3\colon \langle \boldsymbol{e}_j, \boldsymbol{w}\rangle_h = 0\} = \mathcal{S}^1_*(\mathcal{T}_h)^3\,.
\end{aligned}$$

For any $\boldsymbol{q} \in \mathcal{S}^1_*(\mathcal{T}_h)^3$, we compute

$$\begin{aligned}
a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{q}, \boldsymbol{q}) &\overset{(4b)}{=} (1+\alpha^2)\langle (-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{P}_*\boldsymbol{q}, \boldsymbol{q}\rangle_h - \alpha\ell_{\mathrm{ex}}^2\theta k\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{q}), \boldsymbol{q}\rangle_h\,, \\
&\overset{(34),(4d)}{=} (1+\alpha^2)\langle (-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}, -\boldsymbol{\Delta}_h(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\rangle_h + \alpha\ell_{\mathrm{ex}}^2\theta k \left\|\boldsymbol{m}_h^\ell \times \boldsymbol{q}\right\|_h^2 \\
&\overset{(12)}{=} (1+\alpha^2)\left\|\nabla(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\right\|_{\boldsymbol{L}^2(\Omega)}^2 + \alpha\ell_{\mathrm{ex}}^2\theta k \left\|\boldsymbol{m}_h^\ell \times \boldsymbol{q}\right\|_h^2 \\
&\gtrsim h^2 \left\|\boldsymbol{q}\right\|_h^2 + \alpha\ell_{\mathrm{ex}}^2\theta k \left\|\boldsymbol{m}_h^\ell \times \boldsymbol{q}\right\|_h^2 \geq h^2 \left\|\boldsymbol{q}\right\|_h^2\,,
\end{aligned}$$

where the second to last estimate is an inverse estimate on $\mathcal{S}^1_*(\mathcal{T}_h)^3$ derived from the classical inverse estimate on $\mathcal{S}^1(\mathcal{T}_h)^3$ via

$$\begin{aligned}
\|\boldsymbol{q}\|_h^2 = \langle \boldsymbol{q}, \boldsymbol{q}\rangle_h &\overset{(34)}{=} \langle \boldsymbol{q}, -\boldsymbol{\Delta}_h(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\rangle_h \overset{(12)}{=} \langle \nabla\boldsymbol{q}, \nabla(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\rangle_{\boldsymbol{L}^2(\Omega)} \\
&\leq \|\nabla\boldsymbol{q}\|_{\boldsymbol{L}^2(\Omega)} \left\|\nabla(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\right\|_{\boldsymbol{L}^2(\Omega)} \lesssim h^{-1}\|\boldsymbol{q}\|_h \left\|\nabla(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\right\|_{\boldsymbol{L}^2(\Omega)}\,.
\end{aligned}$$

Hence, $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ is coercive on $\bigcap_{\boldsymbol{\lambda}\in\mathbb{R}^3} \mathcal{N}(b_{\mathrm{sp}}(\cdot, \boldsymbol{\lambda})) = \mathcal{S}^1_*(\mathcal{T}_h)^3$ with ellipticity constant proportional to $h^2 > 0$.

- **Step 5:** Unique solvability and reconstruction of $\boldsymbol{v}_h^\ell$.

Clearly, $b_{\mathrm{sp}}\colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3 \to \mathbb{R}$ satisfies the inf-sup condition with constant $(1+\alpha^2) > 0$. Now unique solvability of the saddle point formulation (35a)–(35b) follows from the Brezzi theorem [12, Theorem 4.2.1]. Ultimately, with $(\boldsymbol{q}, \boldsymbol{\lambda}) \in \mathcal{S}^1_*(\mathcal{T}_h)^3 \times \mathbb{R}^3$ denoting the unique solution of (35a)–(35b), the original unknown solution to (30) is reconstructed via $\boldsymbol{v}_h^\ell = (-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q} \oplus \boldsymbol{\lambda} \in \mathcal{S}^1_*(\mathcal{T}_h)^3 \oplus \mathbb{R}^3 = \mathcal{S}^1(\mathcal{T}_h)^3$ and is therefore also unique. □

Remark 4.6. (i) In the third step of the proof of Theorem 4.5, we introduced the unknown $\boldsymbol{q} := -\boldsymbol{\Delta}_h \boldsymbol{v}_* \in \mathcal{S}^1_*(\mathcal{T}_h)^3$. This idea is inspired by [45, Section 2.3], where the authors subsequently use the Browder–Minty lemma for monotone operators to prove well-posedness of their proposed finite difference LLG integrator based on the second-order backward differentiation formula.
(ii) In Step 4 of the proof of Theorem 4.5, as the new unknown $\boldsymbol{q} = -\boldsymbol{\Delta}_h \boldsymbol{v}_h^\ell$ comprises second-order derivatives of the original unknown, it is not surprising that the ellipticity constant for the bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ scales proportionally to $h^2 > 0$.
(iii) Since $\mathcal{M}_h \subset \mathcal{S}^1(\mathcal{T}_h)^3$ and the predictors of Algorithm 3.1 and Algorithm 4.1 coincide, the proof of Theorem 4.5 is not only an alternative proof to Theorem 4.3, but also to Theorem 3.3, which additionally extends both theorems to the critical value $\alpha = 0$.
(iv) Consequently, Algorithm 3.1 is not only a mass-lumped version of the tangent plane scheme [3], but additionally it is well-posed for the Schrödinger map equation ($\alpha = 0$).
(v) Even though the predictor of Algorithm 3.1 written in the form (17) coincides with the predictor of the tangent plane scheme up to the used integration rule, well-posedness of the tangent plane scheme for the limit case $\alpha = 0$ remains open. Indeed, the proof of Theorem 4.5 relies heavily on mass-lumped integration, and we did not succeed to transfer the proof to exact integration used in the original tangent plane scheme.

4.4. Including the lower-order contributions

We consider the case when the effective field comprises linear lower-order energy contributions $\boldsymbol{\pi}(\boldsymbol{m})$ such as, in particular, the nonlocal stray field $\boldsymbol{h}_{\mathrm{s}}$, i.e., $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$. Then the predictor step of the original second-order integrator proposed in [31, Algorithm 2] is identical to (21), i.e., lower-order terms are treated implicitly in time. Due to the nonlocality of the stray field this is unattractive in practice as described in Section 3.3. Hence, analogously to Section 3.3, we aim to treat the lower-order terms $\boldsymbol{\pi}(\boldsymbol{m})$ explicitly in time. However, to avoid spoiling the scheme's potential second-order accuracy in time, which was observed experimentally in [31], the modification is slightly more involved:

In Section 3.3 an error of order $\mathcal{O}(k)$ is introduced to the system (21) by approximating $\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) \approx \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell)$. Since Algorithm 3.1 is a first-order scheme, this modification did not deteriorate the order of convergence of the algorithm.

To preserve the potential second-order of Algorithm 4.1, we use a higher-order approximation to $\boldsymbol{\pi}(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell)$: Recall, that $\boldsymbol{\pi}$ is a linear operator and that $\boldsymbol{v}_h^\ell$ is an approximation of $\partial_t \boldsymbol{m}(t_\ell)$. Motivated by the Taylor expansion $\boldsymbol{m}(t_\ell) = \boldsymbol{m}(t_{\ell-1}) + k\partial_t \boldsymbol{m}(t_\ell) + \mathcal{O}(k^2)$, and hence $\boldsymbol{m}(t_\ell) + \theta k \partial_t \boldsymbol{m}(t_\ell) = (1+\theta)\boldsymbol{m}(t_\ell) - \theta \boldsymbol{m}(t_{\ell-1}) + \mathcal{O}(k^2)$, we introduce a second-order error $\mathcal{O}(k^2)$ to the system (21) via the approximation

$$\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) \approx (1+\theta)\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta \boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1})\,. \tag{36}$$

Only the leading-order exchange contribution is treated implicitly in time, while the lower-order contributions are treated explicitly. Due to the higher-order approximation of $\boldsymbol{\pi}_h(\boldsymbol{v}_h^\ell)$, this does not spoil the observed second-order of the scheme and it is computationally much more attractive. To sum up, we consider the following algorithm.

Algorithm 4.7 (`PC2+IMEX`). ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Preprocessing: *Compute* $\boldsymbol{m}_h^1 \in \mathcal{M}_h$, *e.g., by Algorithm 4.1.*
Loop: *For all time-steps* $\ell = 1, \dots, L-1$, *iterate:*

(i) *Compute* $\mathbb{P}_h((1+\theta)\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta\boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1}) + \boldsymbol{f}^{\ell+\theta}) \in \mathcal{S}^1(\mathcal{T}_h)^3$.
(ii) *Compute* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$\begin{aligned}
&(1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h\rangle_h \\
&= -\langle \boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h((1+\theta)\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta\boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1}) + \boldsymbol{f}^{\ell+\theta})], \boldsymbol{w}_h\rangle_h \\
&\quad -\alpha\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h((1+\theta)\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta\boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1}) + \boldsymbol{f}^{\ell+\theta})]), \boldsymbol{w}_h\rangle_h.
\end{aligned} \tag{37}$$

(iii) *Compute* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$(1+\alpha^2)\langle d_t \boldsymbol{m}_h^{\ell+1}, \boldsymbol{w}_h\rangle_h \tag{38}$$

$$
\begin{aligned}
&= -\langle \boldsymbol{m}_h^{\ell+1/2} \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+1/2})], \boldsymbol{w}_h\rangle_h \\
&\quad - \alpha \langle \boldsymbol{m}_h^{\ell+1/2} \times \big((\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) \\
&\qquad \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+1/2})]\big), \boldsymbol{w}_h\rangle_h.
\end{aligned}
$$

Output: *Sequence of discrete functions* $\big\{(\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1})\big\}_{\ell=0}^{L-1}$.

Remark 4.8. (i) In the preprocessing step of Algorithm 4.7 also other integrators may be used to compute $\boldsymbol{m}_h^1 \in \mathcal{M}_h$. As long as the approximation $\boldsymbol{m}_h^1$ is second-order accurate, the potential second-order accuracy of Algorithm 4.1 is preserved by Algorithm 4.7. (Note that first-order accurate integrators usually only introduce a quadratic error per time-step.)
(ii) Algorithm 4.7 is also well-posed in practice, when effects of inexact (iterative) solvers are accounted for, i.e., (37) is unconditionally well-posed for arbitrary $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3 \supsetneq \mathcal{M}_h$. As lower-order terms are treated explicitly in time, proving well-posedness follows the lines of the proof of Theorem 4.5 with adjusted linear form $F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell] \rightsquigarrow F_{\mathrm{imex}}[\boldsymbol{m}_h^\ell, \boldsymbol{m}_h^{\ell-1}]$.
(iii) In addition to the evaluation of $\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell)$ in the predictor (37) of Algorithm 4.7, another evaluation $\boldsymbol{\pi}_h(\boldsymbol{v}_h^\ell)$ is required in the corrector (38). This second evaluation of $\boldsymbol{\pi}_h$ per time-step can be avoided by applying (36) with $\theta = 1/2$ to the corrector (38), which leads to the second-order approximation $\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) \approx (3/2)\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - (1/2)\boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1})$. Hence, with this modification one arrives at a cheaper version of `PC2+IMEX` requiring only one evaluation of $\boldsymbol{\pi}_h$ per time-step.

5. Numerical experiments

This section provides some numerical experiments for Algorithm 3.1 and Algorithm 4.1 from [31], as well as their respective IMEX versions proposed in this work, namely Algorithm 3.5 and Algorithm 4.7, respectively. In Section 5.1 we verify the correctness of the proposed integrators (`PC1+IMEX` and `PC2+IMEX`) on the benchmark problem μMAG #4 from [36]. In Section 5.2 the experimental rates of Algorithm 3.1 (`PC1`) and Algorithm 4.1 (`PC2`) reported in [31] are confirmed. Moreover, the experiment shows that lower-order terms can appropriately be treated explicitly in time by Algorithm 3.5 (`PC1+IMEX`) and Algorithm 4.7 (`PC2+IMEX`), respectively, without spoiling the rate of convergence.

All computations have been performed with our micromagnetic software module Commics [37], based on the open-source finite element library Netgen/NGSolve [40]. In Commics, the stray field $\boldsymbol{h}_{\mathrm{s}}$ is computed via the hybrid FEM-BEM approach from [25]; see also [38, Algorithm 12]. We note that meshes generated by Netgen in general do not satisfy the angle condition (24). All experiments were repeated on structured meshes satisfying the angle condition leading to the same results (not displayed).

5.1. μMAG standard problem #4

We verify the practical applicability of the proposed integrators `PC1+IMEX` and `PC2+IMEX` (we choose $\theta = 1/2$) by computing a physically relevant example. To this end, we consider μMAG standard problem #4 [36], which simulates the switching of the magnetization in a thin permalloy layer.

The objective is the simulation of the magnetization dynamics in a thin permalloy film of dimensions 500 nm $\times$ 125 nm $\times$ 3 nm under the influence of a constant applied external field. The involved physical constants and material parameters are the gyromagnetic ratio $\gamma_0 = 2.211 \cdot 10^5$ m/C, the permeability of vacuum $\mu_0 = 4\pi \cdot 10^{-7}$ N/A^2, the saturation magnetization $M_{\mathrm{s}} = 8.0 \cdot 10^5$ A/m, the exchange stiffness constant $A = 1.3 \cdot 10^{-11}$ J/m, and the Gilbert damping constant $\alpha = 0.02$. Starting from a so-called equilibrium S-state [36], the experiment consists in applying the constant applied field $\mu_0 \boldsymbol{H}_{\mathrm{ext}} = (-24.6, 4.3, 0)$ mT for 3 ns.

For the rescaled form (5) of LLG, the above physical quantities lead to the parameters $\ell_{\mathrm{ex}} = \sqrt{2A/(\mu_0 M_{\mathrm{s}}^2)}$, $T = 3 \cdot 10^{-9}\gamma_0 M_{\mathrm{s}}$, and $\boldsymbol{f} = \boldsymbol{H}_{\mathrm{ext}}/M_{\mathrm{s}}$, while $\boldsymbol{\pi}(\boldsymbol{m})$ includes only the stray field $\boldsymbol{h}_{\mathrm{s}}$. For the space discretization, we consider a tetrahedral partition of the thin film generated by Netgen [40] into cells of prescribed mesh size 3 nm. This corresponds to 48 796 elements and 16 683 vertices. For the time discretization, we consider a constant physical time-step size of $\Delta t = 0.1$ *ps*, which is connected to the rescaled time-step size k via the relation $k = \gamma_0 M_{\mathrm{s}} \Delta t$.

For comparison, the desired output of this benchmark problem is the evolution of the x-, y- and z-component of the spatially averaged magnetization. Fig. 1 shows, that our results match those computed by the finite difference code OOMMF [42] available on the μMAG homepage [36].

5.2. Empirical convergence rates for LLG

We aim to illustrate the accuracy and the computational effort of the following four algorithms:

- `PC1`: fully implicit first-order scheme proposed in [31] and recalled in Algorithm 3.1;
- `PC1+IMEX`: `PC1` with explicit treatment of the lower-order terms as proposed in this work and formulated in Algorithm 3.5;

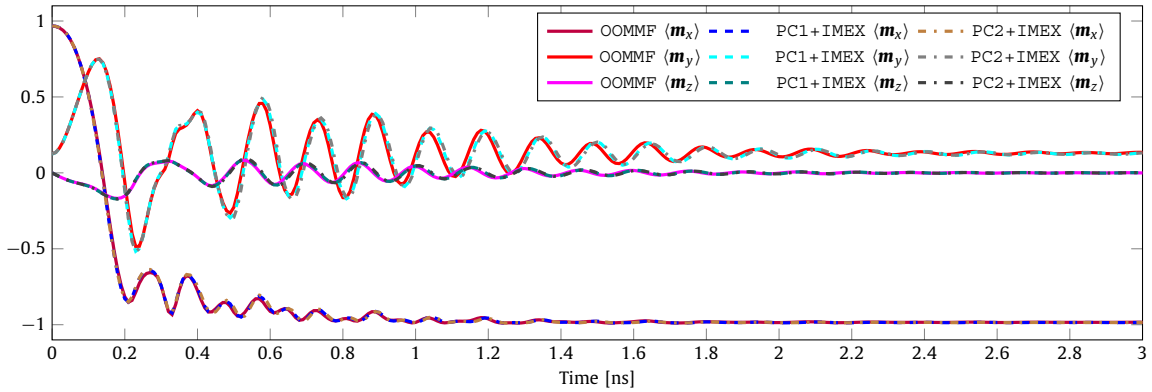


Fig. 1. μMAG standard problem #4 from Section 5.1: Time evolution of the spatially averaged magnetization components computed with Algorithm 3.5 (PC1+IMEX) and Algorithm 4.7 (PC2+IMEX) compared to the results of OOMMF.

- PC2: fully implicit second-order scheme proposed in [31] and recalled in Algorithm 4.1;
- PC2+IMEX: PC2 with explicit treatment of the lower-order terms as proposed in this work and formulated in Algorithm 4.7;

For all integrators we choose $\theta = 1/2$. To obtain experimental convergence rates in time, we use the model problem proposed in [38]: We consider the initial boundary value problem (5) with $\Omega = (0,1)^3$, $\boldsymbol{m}^0 \equiv (1,0,0)$, $\alpha = 1$, and $T = 5$. For the effective field (5d), we choose $\ell_{\mathrm{ex}} = 1$, a constant applied field $\boldsymbol{f} \equiv (-2,-0.5,0)$, as well as an operator $\boldsymbol{\pi}$ which consists only of the stray field, i.e., $\boldsymbol{\pi}(\boldsymbol{m}) = \boldsymbol{h}_{\mathrm{s}}(\boldsymbol{m})$.

For the predictor step in PC1 and PC2, respectively, we solve (21). Since $\boldsymbol{\pi}_h$ effectively depends on $\boldsymbol{v}_h^\ell$ in (21), the linear system in the predictor step of Algorithm 3.1 and Algorithm 4.1 is solved with an inner fixed-point iteration which is stopped as soon as an accuracy of 10^{-10} (of $\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}$) is reached. Other arising linear systems are solved with the generalized minimal residual method (or with the conjugate gradient method for the hybrid FEM-BEM approach) with tolerance 10^{-12}. For the spatial discretization we consider a fixed triangulation $\mathcal{T}_h$ of Ω generated by Netgen, which consists of 3939 elements and 917 nodes (prescribed mesh size $h = 1/8$).

Since the exact solution of the problem is unknown, to compute the empirical convergence rates, we consider a reference solution $\boldsymbol{m}_{h,k_{\mathrm{ref}}}$ computed with the IMEX version of the second-order midpoint scheme from [38] using the above mesh and the time-step size $k_{\mathrm{ref}} = 2 \cdot 10^{-4}$.

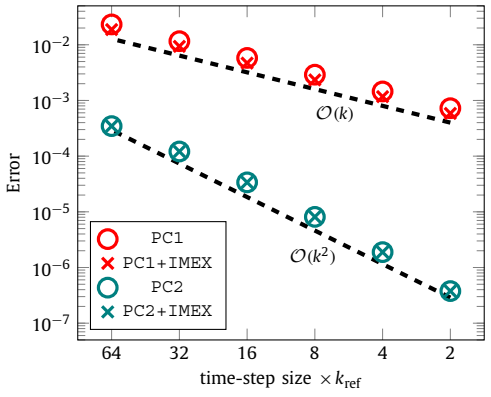
Fig. 2(a) visualizes the experimental order of convergence of the four integrators. As expected, PC2 and PC2+IMEX lead to second-order convergence in time. Essentially, both integrators even lead quantitatively to the same accuracy of the numerical solution. PC1 as well as PC1+IMEX yield first-order convergence. Differently from the classical θ-method for linear second-order parabolic PDEs, due to the tangent plane constraint and the presence of the nodal projection, the PC1 integrator with $\theta = 1/2$ (Crank–Nicolson-type) does not lead to any improvement of the convergence order in time (from first-order to second-order); see [6] for a formal analysis in the case of the tangent plane scheme.

In Fig. 2(b), we plot the cumulative computational costs for the integration up to the final time T. The computational effort improves considerably if the lower-order terms (i.e., the stray field) are integrated explicitly in time, since then the costly inner fixed-point iteration to solve (21) is omitted. Due to the more sophisticated corrector step in Algorithm 4.1 and Algorithm 4.7, the second-order schemes PC2 and PC2+IMEX are (slightly) more costly than their first-order counterparts PC1 and PC1+IMEX, respectively.

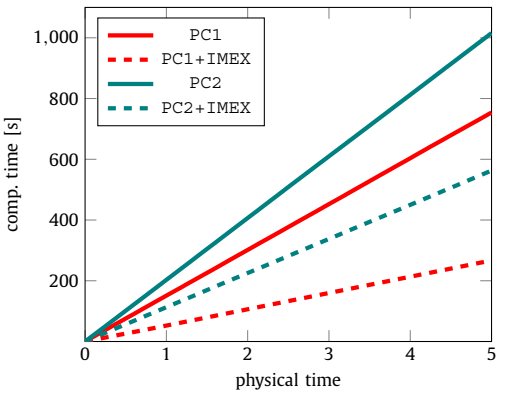
Further, we repeat the experiment for different values of $\theta \in [0,1]$ for both, PC1+IMEX and PC2+IMEX. The results for PC1+IMEX in Fig. 3(a) confirm that the strong CFL condition $k = o(h^2)$, which is imposed to obtain stability and convergence of PC1+IMEX (see Remark 3.7(ii)) with $\theta < 1/2$, are also crucial in practice. As expected, the observed order of convergence of PC1+IMEX is unaffected by the choice of $\theta \in [0,1]$.

The results for PC2+IMEX shown in Fig. 3(b) are quite surprising: While for $\theta \neq 1/2$, the simulation is not stable for larger time-step sizes $k > 0$, still second-order convergence is observed for all $0 \leq \theta \leq 1$ as the time-step size k decreases below a certain threshold. The preserved second-order accuracy for $\theta \neq 1/2$ might be a consequence of the degree of implicitness θ only appearing in the predictor, but not in the corrector of the scheme. In contrast to stability for PC1+IMEX, the results of this experiment indicate that for stability of PC2+IMEX more restrictive CFL conditions are necessary for $\theta \neq 1/2$ than for $\theta = 1/2$. This observation is further investigated in Section 5.3.

Overall, the proposed PC2+IMEX integrator with $\theta = 1/2$ appears to be the method of choice with respect to experimental stability, computational time, and empirical accuracy.

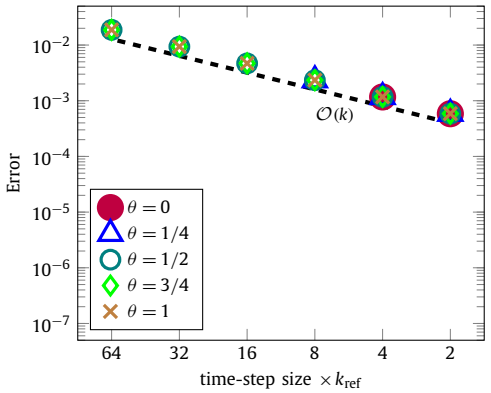


(a) Error $\max_{\ell=0,\dots,L} \|\boldsymbol{m}_{h,k_{\text{ref}}}^{\ell} - \boldsymbol{m}_{h,k}^{\ell}\|_{\boldsymbol{H}^1(\Omega)}$ for $k = 2^{\ell} k_{\text{ref}}$ with $\ell \in \{1,2,3,4,5,6\}$ and $k_{\text{ref}} = 2 \cdot 10^{-4}$.

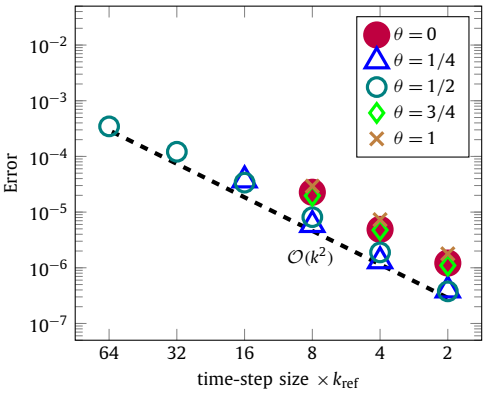


(b) Cumulative computational time for $k = 8 \cdot 10^{-4}$. Costs improve considerably for the IMEX versions.

Fig. 2. Experiments of Section 5.2: Order of convergence (left) and cumulative computational time (right) of the integrators for $\theta = 1/2$.



(a) Recomputation of Fig. 2(a) for PC1+IMEX with various $\theta \in [0,1]$.



(b) Recomputation of Fig. 2(a) for PC2+IMEX with various $\theta \in [0,1]$.

Fig. 3. Experiments of Section 5.2: Order of convergence and stability for PC1+IMEX and PC2+IMEX for different values of $\theta \in [0,1]$. Stability is lost for PC1+IMEX (left) with $\theta = 0$ for $k \geq 8 \cdot k_{\text{ref}}$, and with $\theta = 1/4$ for $k \geq 16 \cdot k_{\text{ref}}$; for PC2+IMEX (right) with $\theta \in \{0, 3/4, 1\}$ for $k \geq 16 \cdot k_{\text{ref}}$, and with $\theta = 1/4$ for $k \geq 32 \cdot k_{\text{ref}}$.

5.3. Experimental stability of PC2

We demonstrated the potential of (the IMEX version of) the second-order predictor-corrector scheme PC2 (PC2+IMEX) in Section 5.1 and Section 5.2. Our analysis guarantees unconditional well-posedness of the proposed second-order integrators in theory (Theorem 4.3) and in practice (Theorem 4.5). However, neither the present work nor [31] include a rigorous analysis on the stability of the second-order predictor-corrector scheme PC2 (Algorithm 4.1), or its variant PC2+IMEX (Algorithm 4.7). More precisely, it remains unclear whether the prescription of a CFL condition $k = o(h^{\beta})$ for some $\beta > 0$ is sufficient to prove a discrete energy estimate of the form

$$\left\| \nabla \boldsymbol{m}_h^J \right\|^2 \leq \left\| \nabla \boldsymbol{m}_h^0 \right\|^2 \qquad \text{for all } J = 0, \dots, L, \tag{39}$$

where we omitted any lower-order contributions; see, e.g., (27) for the full discrete energy estimate for PC1+IMEX.

Hence, we close this section by a numerical study investigating the stability of PC2. Note that PC2+IMEX coincides with PC2 for the exchange only case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \Delta \boldsymbol{m}$ of LLG, which is considered in the following experiments. Motivated by the observations on stability of PC2+IMEX in Fig. 3(b), particular focus is put on the dependence on $0 \leq \theta \leq 1$, which controls the degree of implicitness in the predictor step (30).

5.3.1. Setup

We consider the partition $\mathcal{T}_h$ of the unit cube from Section 5.2. For a non-uniform initial condition $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, we consider the exchange only case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \Delta \boldsymbol{m}$ of LLG and relax the dynamics until the (uniform) equilibrium state is

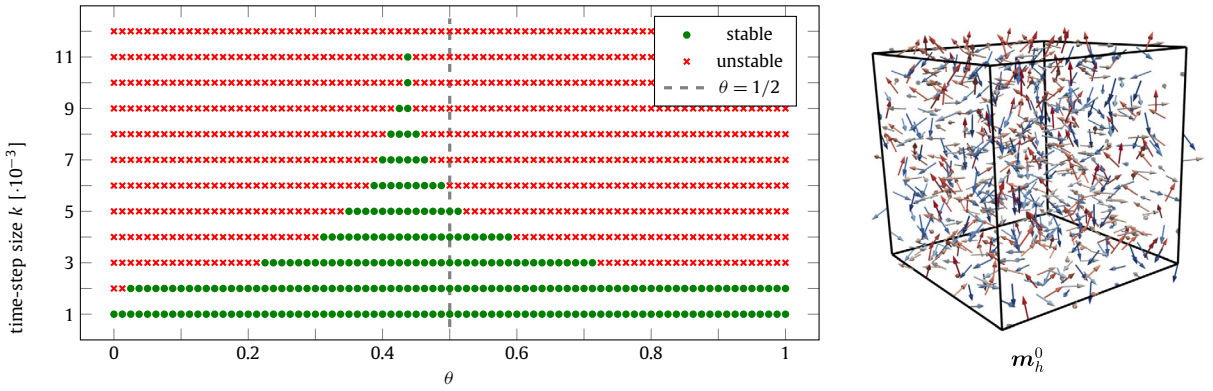


Fig. 4. Experiment of Section 5.3.2. Right: Random state $\boldsymbol{m}_h^0$ colored by the z-component; red pointing upwards, blue downwards. Left: For all $\theta = 0/80, 1/80, \dots, 80/80$ and all $k = 1 \cdot 10^{-3}, 2 \cdot 10^{-3}, \dots, 12 \cdot 10^{-3}$, the stability of PC2 is investigated.

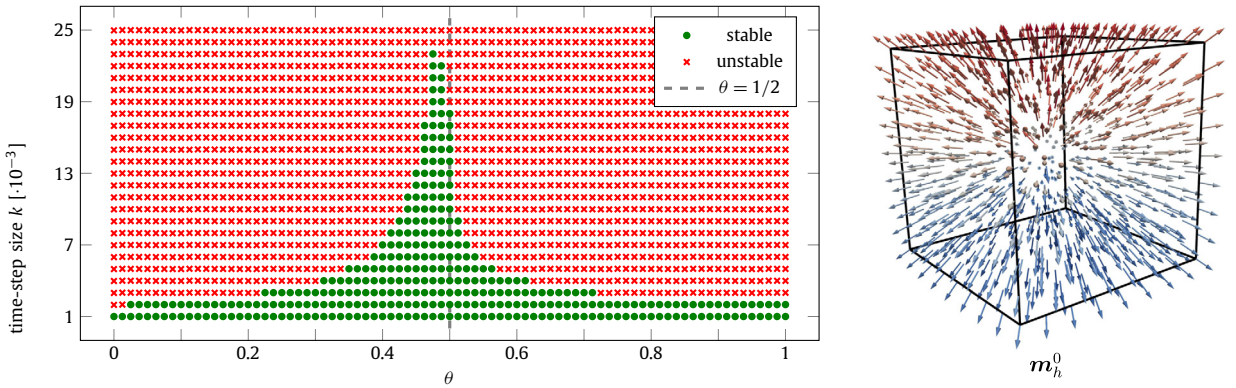


Fig. 5. Experiment of Section 5.3.3: Right: Hedgehog state $\boldsymbol{m}_h^0$ colored by the z-component; red pointing upwards, blue downwards. Left: For all $\theta = 0/80, 1/80, \dots, 80/80$ and all $k = 1 \cdot 10^{-3}, 2 \cdot 10^{-3}, \dots, 25 \cdot 10^{-3}$, the stability of PC2 is investigated.

reached. Due to the absence of any lower-order contributions ($\boldsymbol{\pi} \equiv \boldsymbol{0}$, $\boldsymbol{f} \equiv \boldsymbol{0}$), the equilibrium state is a uniform magnetization in space, and the simulation is successfully stopped as soon as $\|\nabla \boldsymbol{m}_h^L\|^2 \leq 10^{-8}$ for some $L > 0$. If $\|\nabla \boldsymbol{m}_h^{\ell+1}\|^2 \leq \|\nabla \boldsymbol{m}_h^{\ell}\|^2$ for all $\ell = 0, \dots, L-1$, the simulation is considered to be stable for the triangulation $\mathcal{T}_h$ with fixed time-step size $k > 0$ and initial condition $\boldsymbol{m}_h^0 \in \mathcal{M}_h$. If for some $\ell \geq 0$ the energy increases, i.e., if there holds $\|\nabla \boldsymbol{m}_h^{\ell+1}\|^2 > \|\nabla \boldsymbol{m}_h^{\ell}\|^2$, then we abort the simulation and we consider the simulation to be unstable for this combination of $\mathcal{T}_h$, $k > 0$, and $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.

5.3.2. Random initial state

We choose the initial state $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ such that $\{\boldsymbol{m}_h^0(\boldsymbol{z})\}_{\boldsymbol{z} \in \mathcal{N}_h}$ is distributed randomly on $\mathbb{S}^2$.

Fig. 4 shows, that for any fixed $0 \leq \theta \leq 1$ the simulation is stable if the time-step size $k > 0$ is chosen small enough. Clearly, stability of the simulation does not only depend on the chosen time-step size $k > 0$, but also on the parameter θ: Values of θ close to $1/2$ (best at 0.4375 in this experiment) appear to be far less restrictive for the time-step size $k > 0$ than values farther from $1/2$. We note that we repeated this experiment for various random initial states, all producing essentially the same result (not displayed).

5.3.3. Hedgehog state

We repeat the experiment from Section 5.3.2 for $\boldsymbol{m}_h^0$ being the so-called hedgehog state, i.e., considering the cube to be centered around the origin, for each vertex $\boldsymbol{z} \in \mathcal{N}_h$ we set the initial value $\boldsymbol{m}_h^0(\boldsymbol{z}) := \boldsymbol{z}/|\boldsymbol{z}| \in \mathbb{S}^2$.

Fig. 5 shows, that again for any $0 \leq \theta \leq 1$ the simulation is stable if the time-step size $k > 0$ is chosen small enough. As in Section 5.3.2, values of θ close to $1/2$ appear to be far less restrictive for the time-step size $k > 0$ than values farther from $1/2$, with the optimal choice this time closer to $1/2$, precisely at $\theta = 0.475$. Interestingly, for the parameter $\theta \in [0, 1]$ chosen far from $1/2$, the results quantitatively match with those for the random initial state from Section 5.3.2. Closer to $1/2$, however, much larger time-step sizes $k > 0$ allow for stable simulations as for the random initial state.

5.3.4. Variation of the Gilbert damping parameter

We repeat the experiment from Section 5.3.2 for different values of $\alpha = 1/2, 1/4, 1/8, 1/16$.

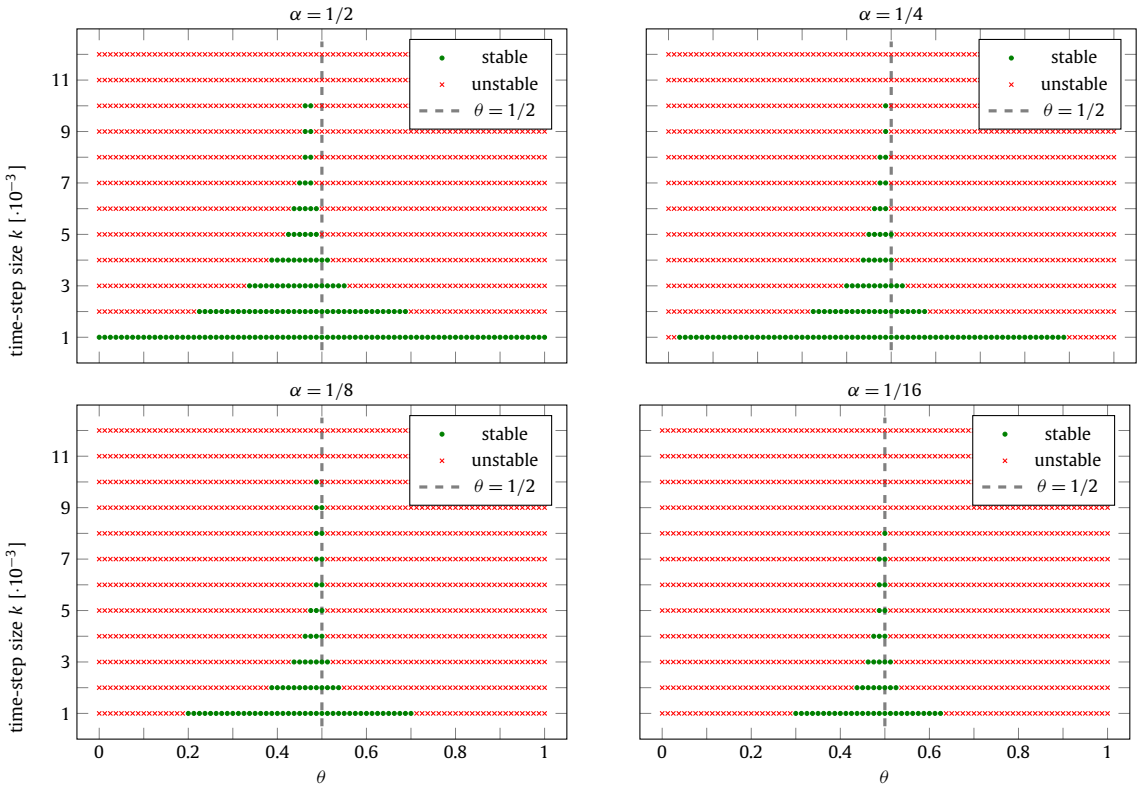


Fig. 6. Experiment of Section 5.3.4: With $\boldsymbol{m}_h^0$ the random state from Fig. 4(right) and different damping parameters $\alpha = 1/2, 1/4, 1/8, 1/16$, for all $\theta = 0/80, 1/80, \dots, 80/80$ and all $k = 1 \cdot 10^{-3}, 2 \cdot 10^{-3}, \dots, 12 \cdot 10^{-3}$, the stability of `PC2 (+IMEX)` is investigated.

Fig. 6 shows that, if the damping parameter α decreases, smaller time-step sizes $k > 0$ are necessary to obtain stable simulations with `PC2 (+IMEX)`. This observation is in agreement with the role played by α in the model, i.e., incorporating dissipation. Again, as previously observed for $\alpha = 1$, values of θ close to $1/2$ allow for larger time-step sizes $k > 0$ than values farther from $1/2$; with the least restrictive choices at $\theta = 0.4375$ for $\alpha = 1$, $\theta = 0.4625$ to 0.475 for $\alpha = 1/2$, $\theta = 0.4875$ for $\alpha = 1/4$, $\theta = 0.4875$ for $\theta = 1/8$, and $\theta = 0.5$ for $\alpha = 1/16$. We obtain analogous results when varying α for the initial hedgehog state (not displayed).

5.3.5. Concluding remarks on the stability of the second-order scheme

All experiments in this section show that, in contrast to `PC1` (Theorem 3.6), larger values of θ do not improve stability of the second-order scheme `PC2`. On the contrary, it is even the case that large values of θ perform as bad as small values of θ. For a generic simulation with `PC2 (+IMEX)`, we suggest to pick the degree of implicitness $\theta = 1/2$ in the predictor. Although, when considering one particular simulation setup, there might be better choices allowing for even larger time-step sizes, the choice $\theta = 1/2$ performed reliably throughout all experiments. In particular, the results from Section 5.3.4 indicate that the deterioration of the optimal θ (with respect to stability) from $1/2$ might occur specifically for large values of α, and quickly vanish as the damping parameter α decreases. Moreover in future works, proving stability of `PC2` under some CFL condition for the special case $\theta = 1/2$ might be a possible first step in theoretically understanding stability of `PC2`. This seems reasonable, as in this special case only the same highest-order term $\boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)$ appears in the predictor and the corrector of `PC2`. Hence these terms partially cancel out, when subtracting the two equations (30)–(31) from each other.

6. Conclusion

In this paper, we have improved the theoretical understanding of the predictor-corrector methods proposed in [31] by establishing unconditional well-posedness of both `PC1` and `PC2` (in Theorem 3.3 and Theorem 4.3, respectively), i.e., we have proved that for each time-step, the arising variational problems admit unique solutions, which was left open in the original paper. By closing this fundamental gap, we have shown that `PC1` is not only closely related to the first-order tangent plane scheme of [3,14], but actually can even be interpreted as a slight modification of it; see Remark 3.4(iii). Our well-posedness analysis is based on a reformulation of the predictor system, which crucially exploits nodewise unit-length $\boldsymbol{m}_h^\ell \in \boldsymbol{\mathcal{M}}_h$. This property is satisfied by the approximations generated by `PC1` thanks to the use of the nodal projection

in the corrector step. In PC2, the unit-length constraint is imposed by the corrector implicitly via a constraint-preserving variational formulation. However, as linear systems are not solved exactly in practice, one cannot expect the unit-length constraint to hold at machine precision. Hence, strictly speaking, the well-posedness analysis of PC2 established in Theorem 4.3 does not cover a practical implementation of the scheme. To cope with this problem, we establish a decomposition of the finite element space, which does not only allow us to prove unconditional well-posedness of the practical version of PC2 (Theorem 4.5), but also to extend the result, for both PC1 and PC2 (theoretical and practical), to the limit case $\alpha = 0$ (Schrödinger map equation); see Remark 3.4(iv), Remark 4.4(ii), and Remark 4.6(iii)–(iv).

Furthermore, following [14,38,21], we propose implicit-explicit versions of PC1 and PC2. When considering magnetization dynamics involving the full effective field — more precisely, dynamics including the nonlocal stray field — the proposed IMEX versions PC1+IMEX and PC2+IMEX are computationally much more attractive than their original counterparts. Experimentally, we demonstrate the applicability of the IMEX schemes to a physically relevant benchmark problem in Section 5.1, and their improved efficiency in Section 5.2. The experiments in Section 5.2 also validate their preservation of the experimental first- and second-order accuracies of PC1 and PC2, respectively.

While from a practical point of view the original integrators PC1 and PC2 from [31] are computationally unattractive due to their implicit integration of the nonlocal stray field contribution, effectiveness of the IMEX integrators PC1+IMEX and PC2+IMEX proposed in this work is comparable to state-of-the-art schemes from the literature: PC1+IMEX is experimentally first-order accurate in time and requires one stray field evaluation in addition to the solution of one sparse linear system per time-step, which makes it directly comparable to the first-order tangent plane scheme of [3,14]. PC2+IMEX is second-order accurate in time and requires one stray field evaluation (see Remark 4.8(iii)) in addition to the solution of two sparse linear systems per time-step. Hence, it is slightly more expensive than, but comparable to, the second-order implicit-explicit tangent plane scheme from [21], which requires one stray field evaluation in addition to the solution of only one sparse linear system per time-step. At the same time, PC2+IMEX is considerably cheaper than the second-order implicit-explicit midpoint scheme from [38], which requires one stray field evaluation in addition to the approximate solution of a nonlinear system per time-step.

Stability of PC1+IMEX is established in Theorem 3.6. As pointed out in Remark 3.7, this leads to convergence results matching those of the first-order tangent plane scheme of [3,14] as well as recovering those for the original integrator PC1 in [31], in particular guaranteeing unconditional convergence for the choices $1/2 < \theta \leq 1$. From a theoretical point of view, stability and convergence of PC2(+IMEX), not addressed in [31], remain open also in our analysis and will be the subject of future research. Instead, we shed some light on the stability of PC2(+IMEX) by a numerical study. In particular, our experiments summarized in Section 5.3.5 suggest that investigation of the special case $\theta = 1/2$ might be the most promising first step towards establishing a first rigorous stability result for PC2(+IMEX).

References

[1] C. Abert, G. Hrkac, M. Page, D. Praetorius, M. Ruggeri, D. Suess, Spin-polarized transport in ferromagnetic multilayers: an unconditionally convergent FEM integrator, Comput. Math. Appl. 68 (6) (2014) 639–654.
[2] G. Akrivis, M. Feischl, B. Kovács, C. Lubich, Higher-order linearly implicit full discretization of the Landau-Lifshitz-Gilbert equation, Math. Comput. 90 (329) (2021) 995–1038.
[3] F. Alouges, A new finite element scheme for Landau–Lifchitz equations, Discrete Contin. Dyn. Syst., Ser. S 1 (2) (2008) 187–196.
[4] F. Alouges, P. Jaisson, Convergence of a finite element discretization for the Landau–Lifshitz equation in micromagnetism, Math. Models Methods Appl. Sci. 16 (2) (2006) 299–316.
[5] F. Alouges, A. Soyeur, On global weak solutions for Landau–Lifshitz equations: existence and nonuniqueness, Nonlinear Anal. 18 (11) (1992) 1071–1084.
[6] F. Alouges, E. Kritsikis, J. Steiner, J.-C. Toussaint, A convergent and precise finite element scheme for Landau–Lifschitz–Gilbert equation, Numer. Math. 128 (3) (2014) 407–430.
[7] R. An, Optimal error estimates of linearized Crank-Nicolson Galerkin method for Landau-Lifshitz equation, J. Sci. Comput. 69 (1) (2016) 1–27.
[8] S. Bartels, Stability and convergence of finite-element approximation schemes for harmonic maps, SIAM J. Numer. Anal. 43 (1) (2005) 220–238.
[9] S. Bartels, Numerical Methods for Nonlinear Partial Differential Equations, Springer, Cham, 2015.
[10] S. Bartels, A. Prohl, Convergence of an implicit finite element method for the Landau–Lifshitz–Gilbert equation, SIAM J. Numer. Anal. 44 (4) (2006) 1405–1419.
[11] S. Bartels, J. Ko, A. Prohl, Numerical analysis of an explicit approximation scheme for the Landau–Lifshitz–Gilbert equation, Math. Comput. 77 (262) (2008) 773–788.
[12] D. Boffi, F. Brezzi, M. Fortin, Mixed Finite Element Methods and Applications, Springer, Berlin, 2013.
[13] W.F. Brown, Micromagnetics, Interscience Publishers, New York, 1963.
[14] F. Bruckner, M. Feischl, T. Führer, P. Goldenits, M. Page, D. Praetorius, M. Ruggeri, D. Suess, Multiscale modeling in micromagnetics: existence of solutions and numerical integration, Math. Models Methods Appl. Sci. 24 (13) (2014) 2627–2662.
[15] G. Carbou, P. Fabrie, Regular solutions for Landau-Lifschitz equation in a bounded domain, Differ. Integral Equ. 14 (2) (2001) 213–229.
[16] J. Chen, C. Wang, C. Xie, Convergence analysis of a second-order semi-implicit projection method for Landau-Lifshitz equation, Appl. Numer. Math. 168 (2021) 55–74.
[17] I. Cimrák, A survey on the numerics and computations for the Landau–Lifshitz equation of micromagnetism, Arch. Comput. Methods Eng. 15 (3) (2008) 277–309.
[18] I. Cimrák, Convergence result for the constraint preserving mid-point scheme for micromagnetism, J. Comput. Appl. Math. 228 (1) (2009) 238–246.
[19] M. D'Aquino, C. Serpico, G. Miano, Geometrical integration of Landau–Lifshitz–Gilbert equation based on the mid-point rule, J. Comput. Phys. 209 (2) (2005) 730–753.
[20] G. Di Fratta, M. Innerberger, D. Praetorius, Weak-strong uniqueness for the Landau–Lifshitz–Gilbert equation in micromagnetics, Nonlinear Anal., Real World Appl. 55 (2020) 103122.
[21] G. Di Fratta, C.-M. Pfeiler, D. Praetorius, M. Ruggeri, B. Stiftner, Linear second-order IMEX-type integrator for the (Eddy current) Landau–Lifshitz–Gilbert equation, IMA J. Numer. Anal. 40 (4) (2020) 2802–2838.

[22] E. Dumas, F. Sueur, On the weak solutions to the Maxwell–Landau–Lifshitz equations and to the Hall-Magneto-Hydrodynamic equations, Commun. Math. Phys. 330 (3) (2014) 1179–1225.
[23] M. Feischl, T. Tran, The Eddy Current–LLG equations: FEM-BEM coupling and a priori error estimates, SIAM J. Numer. Anal. 55 (4) (2017) 1786–1819.
[24] M. Feischl, T. Tran, Existence of regular solutions of the Landau-Lifshitz-Gilbert equation in 3D with natural boundary conditions, SIAM J. Math. Anal. 49 (6) (2017) 4470–4490.
[25] D.R. Fredkin, T.R. Koehler, Hybrid method for computing demagnetization fields, IEEE Trans. Magn. 26 (2) (1990) 415–417.
[26] H. Gao, Optimal error estimates of a linearized backward Euler FEM for the Landau-Lifshitz equation, SIAM J. Numer. Anal. 52 (5) (2014) 2574–2593.
[27] C.J. García-Cervera, Numerical micromagnetics: a review, Bol. Soc. Esp. Mat. Apl. SeMA 39 (2007) 103–135.
[28] T.L. Gilbert, A phenomenological theory of damping in ferromagnetic materials, IEEE Trans. Magn. 40 (6) (2004) 3443–3449.
[29] B. Guo, M.-C. Hong, The Landau–Lifshitz equation of the ferromagnetic spin chain and harmonic maps, Calc. Var. Partial Differ. Equ. 1 (3) (1993) 311–334.
[30] E. Kim, K. Lipnikov, The mimetic finite difference method for the Landau-Lifshitz equation, J. Comput. Phys. 328 (2017) 109–130.
[31] E. Kim, J. Wilkening, Convergence of a mass-lumped finite element method for the Landau–Lifshitz equation, Q. Appl. Math. 76 (2018) 383–405.
[32] M. Kruzik, A. Prohl, Recent developments in the modeling, analysis, and numerics of ferromagnetism, SIAM Rev. 48 (3) (2006) 439–483.
[33] L. Landau, E. Lifshitz, On the theory of the dispersion of magnetic permeability in ferromagnetic bodies, Phys. Z. Sowjetunion 8 (1935) 153–168.
[34] F. Lin, C. Wang, The Analysis of Harmonic Maps and Their Heat Flows, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.
[35] C. Melcher, Existence of partially regular solutions for Landau–Lifshitz equations in $\mathbb{R}^3$, Commun. Partial Differ. Equ. 30 (4) (2005) 567–587.
[36] NIST micromagnetic modeling activity group (μMAG) website, https://www.ctcms.nist.gov/~rdm/mumag.org.html. (Accessed 30 March 2022).
[37] C.-M. Pfeiler, M. Ruggeri, B. Stiftner, L. Exl, M. Hochsteger, G. Hrkac, J. Schöberl, N.J. Mauser, D. Praetorius, Computational micromagnetics with commics, Comput. Phys. Commun. 248 (2020) 106965.
[38] D. Praetorius, M. Ruggeri, B. Stiftner, Convergence of an implicit-explicit midpoint scheme for computational micromagnetics, Comput. Math. Appl. 75 (5) (2018) 1719–1738.
[39] A. Prohl, Computational Micromagnetism, Teubner, Wiesbaden, 2001.
[40] J. Schöberl, Netgen/NGSolve, https://ngsolve.org/. (Accessed 30 March 2022).
[41] P.L. Sulem, C. Sulem, C. Bardos, On the continuous limit for a system of classical spins, Commun. Math. Phys. 107 (3) (1986) 431–454.
[42] The Object Oriented MicroMagnetic Framework (OOMMF) project at ITL/NIST, https://math.nist.gov/oommf/. (Accessed 30 March 2022).
[43] A. Visintin, On Landau–Lifshitz' equations for ferromagnetism, Jpn. J. Appl. Math. 2 (1) (1985) 69–84.
[44] X.-P. Wang, C.J. García-Cervera, W. E, A Gauss-Seidel projection method for micromagnetics simulations, J. Comput. Phys. 171 (1) (2001) 357–372.
[45] C. Xie, C.J. García-Cervera, C. Wang, Z. Zhou, J. Chen, Second-order semi-implicit projection methods for micromagnetics simulations, J. Comput. Phys. 404 (2020) 109104.